The e-value and the Full Bayesian Significance Test: Logical Properties and Philosophical Consequences.

Julio Michael Stern*
Carlos Alberto de Bragança Pereira†
Marcelo de Souza Lauretto‡
Luís Gustavo Esteves§
Rafael Izbicki¶
Rafael Bassi Stern‖
Marcio Alves Diniz**
Wagner de Souza Borges††

Dedicated to Walter Alexandre Carnielli, for his 70th birthday

Abstract: This article gives a conceptual review of the e-value, $\mathrm{ev}(H \mid X)$ – the epistemic value of hypothesis H given observations X. This statistical significance measure was developed in order to allow logically coherent and consistent tests of hypotheses, including sharp or precise hypotheses, via the Full Bayesian Significance Test (FBST). Arguments of analysis allow a full characterization of this statistical test by its logical or compositional properties, showing a mutual complementarity between results of mathematical statistics and the logical desiderata lying at the foundations of this theory.

1 Introduction

The e-value, $\mathrm{ev}(H|X)$ – also named the *epistemic-value* of hypothesis H given observations X, or the evidence-value of observations X in favor (or in support) of hypothesis H – is a Bayesian statistical significance measure introduced in 1999 by Carlos Alberto de Bragança Pereira and Julio Michael Stern, together with the FBST – the Full Bayesian Significance Test, see [16]. The definitions of e-value and the FBST were further refined and generalized by subsequent works of several researchers at the University of São Paulo (USP) and the Federal University of São Carlos (UFSCar), in Brazil, including Wagner Borges, Luís Gustavo Esteves, Rafael Izbicki, Regina Madruga, Rafael Bassi Stern, and Sergio Wechsler, see [2, 4, 5, 15, 17].

The e-value was specially designed to assess the *logical truth value* (a.k.a. *statistical significance*) of *sharp* (a.k.a. *precise*) hypotheses in the context of Bayesian statistics. The e-value has desirable

*`jstern@ime.usp.br` , `www.ime.usp.br/~jmstern/`, ORCID: 0000-0003-2720-3871, ResearcherID: C-1128-2013, ID-Lattes: 9582404119292455, Google Scholar: 7spXyx8AAAAJ; IME-USP - Institute of Mathematics and Statistics of the University of Sao Paulo. Rua do Matão, 1010, 05508-900, São Paulo, Brazil.

†`cpereira@ime.usp.br`, ORCID: 0000-0001-6315-7484, ResearcherID: O-5022-2015; IME-USP.

‡`marcelolaureto@usp.br`, ORCID: 0000-0001-5507-2368, ResearcherID: J-1964-2012; EACH-USP - School of Arts, Sciences and Humanities of the University of Sao Paulo. Rua Arlindo Béttio, 1000, 03828-000, São Paulo, Brazil.

§`lesteves@ime.usp.br`, ID-Lattes: 8881931568548501; IME-USP.

¶`rizbicki@ufscar.br`, ORCID: 0000-0003-0379-9690, ResearcherID: J-1160-2014; DS-UFScar - Department of Statistics of the Federal University of Sao Carlos. Rodovia Washington Luis, km 235, 13565-905, São Carlos, Brazil.

‖`rbstern@ufscar.br`, ORCID: 0000-0002-4323-4515, ResearcherID: I-4745-2014; DS-UFSCar.

**`marcio.alves.diniz@gmail.com`, ORCID: 0000-0002-8239-4263, ID-Lattes: 8948404469003829; DS-UFSCar.

††`wborges@ime.usp.br`, ORCID: 0000-0001-5520-5861, ID-Lattes: 0975898314511610; IME-USP.

asymptotic, geometrical (i.e., invariance), and logical (i.e., compositional) properties that allow consistent and coherent evaluation and testing of sharp statistical hypotheses. Furthermore, in applied modeling, the FBST offers an easy-to-implement and powerful statistical test that is fully compliant with Bayesian principles of good inference, like the likelihood principle, see [116, 160, 189, 230].

In the context of statistical test of hypotheses, a *compositional logic* is conveyed by an algebraic formalism that allows the evaluation of truth-functions of composite models and truth-values of composite hypotheses by algebraic operations on the corresponding truth functions of elementary models and truth values of elementary hypotheses. The e-value and the FBST have a rich, expressive, and intuitive compositional logic, while traditional truth-values and accompanying tests offered by either frequentist (classical) statistics, like the p-value, or by Bayesian statistics, like Bayes factors, have important and well-known deficiencies in this regard, specially in cases involving sharp statistical hypotheses. Furthermore, *logically coherent* evaluations and testing of sets and sub-sets of statistical hypotheses should render sequences of inferential reasoning that do not generate internal contradictions or anti-intuitive results. As expected, the e-value and the FBST comply with well-established rules of logical coherence, even in the case of sharp hypotheses, while traditional alternatives often fail to do so.

Asymptotically, $\text{sev}(H|X)$ – the standardized e-value – shares several properties of the p-value, the well-known significance measure of frequentist statistics. This allows the use of the standardized e-value in frequentist-oriented applications, retaining, nevertheless, many theoretical characteristics of the Bayesian framework. Many theoretical developments and practical applications of the e-value and the FBST have already been published in the scientific literature, see [19]. Concise entries about the e-value and the FBST are available at the *International Encyclopedia of Statistical Science* and online at Wiley's *StatsRef*, see [18, 66].

Section 2 reviews the Bayesian statistical framework; Section 3 defines the e-value; Sections 4, 5, and 6 explain the invariance, asymptotic, and compositional properties of the e-value; Section 7 defines the GFBST - the Generalized Full Bayesian Significance Test and its logical properties; Sections 8 and 9 comment on computational implementation and give a detailed numerical example in model selection; Section 10 lists a representative assortment of articles from many practical applications of the e-value and the FBST already published in the scientific literature. Section 11 considers the philosophical consequences of the aforementioned developments by briefly commenting on the *Objective cognitive constructivism* epistemological framework, which was specifically developed to accommodate the formal properties of the e-value and the FBST, and renders a naturalized approach to ontology and metaphysics. Section 12 presents some topics for further research at the interface between Logic and Statistics. One of the objectives of this paper is to foster work in areas of interface or overlap between Logic and Statistics and to stimulate greater cooperation between the two communities. Hence, for the sake of clarity and completeness, this paper includes some basic definitions and explanations in both areas that may be convenient to facilitate mutual understanding of more advanced topics.

2 Bayesian Framework

The basic definition of the e-value is given in the context of a standard parametric model in Bayesian statistics, where observable variables x are generated with a probability density $p(x\,|\,\theta)$. The observable (vectors of) random variables $x \in \mathscr{X} \subseteq \mathbb{R}^s$ belong to the model's *sample space*, while the latent (non-observable) vector $\theta \in \Theta \subseteq \mathbb{R}^t$ belongs to the model's *parameter space*. Statistical inference aims to acquire information about the unknown parameter θ from a sequence of observations, $X = [x^{(1)}; x^{(2)}; \dots x^{(n)}]$, that for notational convenience are stacked in $n \times s$ matrix X.

Since the true value of the parameter, θ^0, is unknown, θ is treated in Bayesian statistics as a (vector) random variable. Available previous information about the parameter is represented by its *a priori* density, $p_0(\theta)$. No previous information about θ is represented by a non-informative *reference density*, $r(\theta)$. From *Bayes rule*, it follows that, after having n independent observations in dataset X, the available information about the parameter θ is represented by the *posterior density*

$$p_n(\theta\,|X) \;=\; (1/c_n)p_0(\theta)\prod_{i=1}^{n} p(x^{(i)}\,|\,\theta) \;=\; (1/c_n)p_0(\theta)p(X\,|\,\theta) \;,$$

where c_n is the appropriate normalization constant.

Whenever possible, we use a relaxed notation leaving implicit the condionalization on the observed dataset, for example, writing $p_n(\theta)$ instead of $p_n(\theta \,|\, X)$. Further details about the Bayesian framework, including appropriate choices for non-informative reference densities, can be found in [31, 123, 160, 173, 176, 235].

A statistical hypothesis H states that the parameter θ^0 generating the observations X belongs to the *hypothesis' set*, Θ_H, a region of Θ constrained by (vector) inequality and equality constraints,

$$\Theta_H = \{\theta \in \Theta \,|\, g(\theta) \leq \mathbf{0} \wedge h(\theta) = \mathbf{0}\} \ .$$

It is common practice in statistics to use a relaxed notation, writing H for both the hypothesis statement and the hypothesis' set.

The presence of q equality constraint makes the hypothesis *sharp* or *precise*, namely, H becomes a proper surface (sub-manifold) of dimension strictly lower than the dimension of the parameter space, that is, $h = \dim(H) = t - q < t = \dim(\Theta)$. In particular, a point hypothesis is a hypothesis of dimension $h = 0$. Regarded as a subset of the t-dimensional parameter space, a sharp hypothesis has volume zero, and should therefore have zero posterior probability, at least for a regular (continuous and differentiable) posterior density, which is usually a natural assumption for statistical modeling, see [11]. These conditions make $\Pr(H)$, the probability of hypothesis H, an inappropriate measure to evaluate sharp hypotheses, unless a special probability measure is created for the hypothesis set, a situation that may be computationally cumbersome and theoretically challenging, see [35, 232]. Nevertheless, sharp hypotheses have a prominent role in science, for the most important statements in exact sciences are *natural laws*. These are formulated as equations in a theory of interest that, in turn, can be expressed by sharp hypotheses in statistical models used to empirically verify the same laws, see [5, 31, 32, 37, 42]. Sharp hypotheses also naturally arise in legal applications, like auditing models for regulatory compliance and concerns related to the burden of proof in related legal cases, see [25, 39, 93, 163, 190, 250, 254, 260]. This state of affairs was the main motivation for defining the e-value as a Bayesian significance measure specially well-suited for sharp hypotheses.

3 The e-value

$\mathrm{ev}\,(H \,|\, X) \in [0,1]$, the e-value, or the epistemic value of hypothesis H given the observed data X, or the evidence given by the observed data X in favor of hypothesis H, and its complement, $\overline{\mathrm{ev}}(H \,|\, X) = 1 - \mathrm{ev}\,(H \,|\, X)$, are defined as follows:

(i) $s(\theta)$, the *surprise function* in a statistical model is defined as the quotient between the posterior and the reference densities in the statistical model, see [15, 35, 139, 150, 202, 212, 213, 251],

$$s(\theta) = p_n(\theta)/r(\theta) \ ;$$

(ii) s^*, the maximum (or supremum) of the surprise function constrained to the hypothesis H, is defined as

$$s^* = \sup_{\theta \in H} s(\theta) \ ,$$

A maximizing argument, $\theta^* \,|\, s^* = s(\theta^*)$, is called a *tangential point*, see Figure 1;

(iii) $T(v)$, the closed lower v-cut of the surprise function, see [35, 139], and its complement, the open upper v-cut of the surprise function, $\overline{T}(v)$, are defined as

$$T(v) = \{\theta \in \Theta \,|\, s(\theta) \leq v\} \ , \quad \overline{T}(v) = \{\theta \in \Theta \,|\, s(\theta) > v\} \ ;$$

The upper v-cut at level $v = s^*$, $\overline{T}(s^*)$, is called the *tangential set*, for its border corresponds to the contour line of the surprise function that is tangential to hypothesis H, see Figure 1.

(iv) $W(v)$, the *truth function* or *Wahrheitsfunktion* at level v, is defined as the posterior probability mass inside the lower v-cut of the surprise function, see [2, 35],

$$W(v) = \int_{T(v)} p_n(\theta)\, d\theta \ ,$$

while its complement is defined as $\overline{W}(v) = 1 - W(v)$;

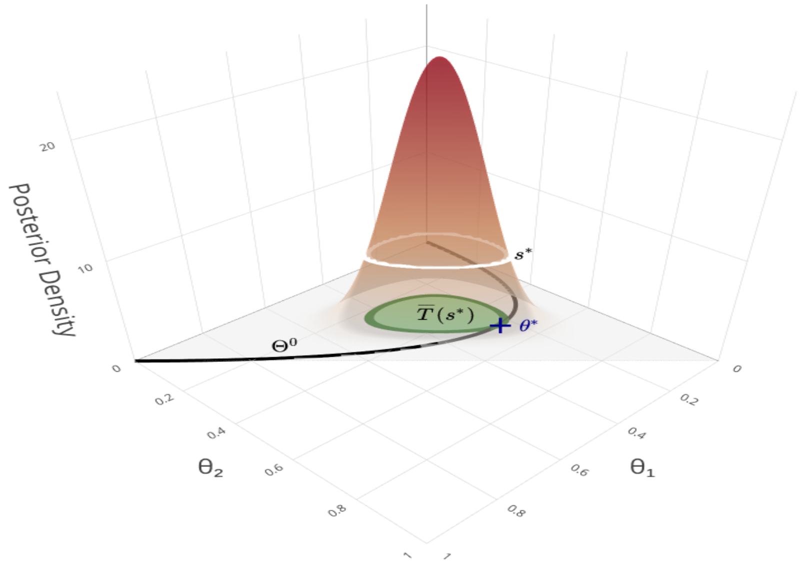

Figure 1: Flat prior, $r(\theta) \propto 1$; Surprise function equal to posterior density, $s(\theta) = p_n(\theta)$; Optimal $\theta^* \in H$; Contour curve at level $s^* = s(\theta^*)$; and Tangential highest probability density set, $\overline{T}(\theta^*)$.

(v) $\text{ev}\,(H\,|\,X)$, the *epistemic value* of hypothesis H given the observed data X, is defined as the truth function $W(v)$ computed at level $v = s^*$, see [2, 16, 17, 35], while its complement, $\overline{\text{ev}}\,(H\,|\,X)$, the evidence given by the observed data X against hypothesis H, has the complementary probability mass,

$$\text{ev}\,(H\,|\,X) = W(s^*)\ , \quad \overline{\text{ev}}\,(H\,|\,X) = \overline{W}(s^*) = 1 - \text{ev}\,(H)\ ;$$

For further details and remarks on the development of the e-value and its historical predecessors, see [16, 25, 35].

In the special case of the flat reference density, $r(\theta) \propto 1$, the surprise function coincides with the posterior probability density function. Its upper cuts, $\overline{T}(v)$, coincide with *highest probability density sets* (HPDS), making it easier to visualize all the elements defined in the last paragraph. Figure 1 depicts a simple example with a unimodal surprise function and a simply connected tangential set, see [16, Sec.4.3]. However, in general, the posterior probability density function may have multiple local maxima, the tangential set may have multiple connected components, and all these components must be taken into account when computing the e-value. In this respect, the e-value differs from traditional uses of HPDS that only consider a single connected component, usually the connected component containing $\widehat{\theta}$, the unconstrained maximum a posteriory estimator, see [35, 123, 140, 184, 208, 231].

4 Invariance and Reference Density

A *reparameterization* of the parameter space, that is, a change in the coordinate system used to map Θ, could "stretch" or "compress" the region surrounding a given location point in the map. This kind of effect is clearly visible comparing geographic maps using different cartographic projections. However, the probability mass inside a given region, as specified by a probability density function, should remain the same, regardless of the coordinate system in use. Hence, the value or "height" of a density function at a given point must change according to the coordinate system in use. This kind of reasoning explains the transformation rules specified by differential and integral calculus for modifying density functions according to transformations of the coordinate system, see [152, 154, 195]. Nevertheless, since the surprise function is defined as the ratio of two densities, $s(\theta) = p_n(\theta)/r(\theta)$, its value remains unchanged by regular (continuous and differentiable) reparamaterizations of the parameter space, and the e-value inherits this important invariance property, see [2, 35].

Invariance is a vital property for good statistical inference for, otherwise, conclusions reached by statistical analysis would depend on the coordinate system being employed, like units of measurement, map projections, forms of data visualization, and other idiosyncrasies or arbitrary choices. Invariance properties of the e-value can be further explained and analyzed in the context of *Information geometry*, where the reference density, $r(\theta)$, can be seen as an implicit representation of the underlying *Information metric* for the parameter space, see [2, 35, 110–112, 115, 117, 145, 157, 160, 173].

5 Asymptotic Consistency

As the number n of observations grows to infinity, that is, in the asymptotic limit $n \to \infty$, we expect a *consistent* statistical procedure to reach the "correct" conclusion. This section analyses the asymptotic properties of the e-value that motivate its use to consistently evaluate a statistical hypothesis. It is easier to describe these asymptotic properties in terms of the *standarized* version of the e-value, defined as follows:

(i) $Q(d,z)$, the *chi-square* cumulative distribution with $d \in \mathbb{N}_+$ degrees of freedom for random variable $z \in [0,\infty]$, is defined by the following analytical expression using the *incomplete gamma function*:

$$Q(d,z) = \frac{\gamma(d/2,z/2)}{\gamma(d/2,\infty)} \ , \quad \gamma(d,z) = \int_0^z y^{d-1} e^{-y} dy \ ;$$

(ii) $\sigma(t,h,c)$, the *standardization function* on arguments $t,h \in \mathbb{N}_+$ and $c \in [0,1]$, is defined by the expression

$$\sigma(t,h,c) = Q\left(t-h, Q^{-1}(t,c)\right) \ ;$$

(iii) $\mathrm{sev}(H\,|\,X)$, the standarized e-value of a hypothesis $H \subset \Theta$ of dimension $h = \dim(H) \leq t = \dim(\Theta)$, is defined as follows:

$$\mathrm{sev}(H\,|\,X) = 1 - \overline{\mathrm{sev}}(H\,|\,X) \ , \quad \overline{\mathrm{sev}}(H\,|\,X) = \sigma(t,h,\overline{\mathrm{ev}}(H\,|\,X)) \ ;$$

The standarized e-value has the following asymptotic properties, under usual continuity and differentiability regularity conditions, see [2]:

(a) If H is true, i.e. $\theta^0 \in H$; and H is slack, i.e. $\dim(H) = \dim(\Theta)$; and θ^0 is in the topological interior of H; Then $\mathrm{ev}\,(H\,|\,X)$ converges (in probability) to 1, as the number of observations increases;

(b) If H is true, i.e. $\theta^0 \in H$; and H is sharp, i.e. $\dim(H) < \dim(\Theta)$; Then $\mathrm{sev}(H\,|\,X)$ converges (in distribution) to $U[0,1]$, the uniform distribution in the unit interval. The standardized e-value shares these asymptotic properties with the p-value – the well-known and widely used significance measure of frequentist statistics. Hence, in a frequentist-oriented application, one could use the e-value as a Bayesian replacement for the p-value, preserving already familiar decision procedures and their interpretations. Moreover, in several modeling applications, the e-value exhibits better convergence characteristics than those of the p-value, see [12, 13, 55, 79, 80]. Further consistency properties of the e-value and higher order asymptotic approximations have been studied and developed in [3, 20–22, 45–47, 94].

6 Compositional Logic of e-values

The e-value and its truth function yield simple algebraic expressions that facilitate the study of a composite model, M, build by serial coupling of k independent statistical models, indexed on $j = 1 \dots k$, $M^{(j)} = \{\Theta^{(j)}, p_0^{(j)}, p_n^{(j)}, r^{(j)}\}$. In this setting, q alternative (or parallel) sets of hypotheses, $H^{(i,j)}$, indexed on $i = 1 \dots q$, are provided for evaluation. This situation is somewhat analogous to the analysis of parallel-serial systems in reliability theory, see [114, 122, 179] and Figure 2. Let us first consider a pure serial system, consisting of k individual models, $M^{(j)}, j = 1 \dots k$, each one contemplating a single hypothesis, $H^{(1,j)}$.

Since the individual models are independent, the joint posterior density, reference, and surprise function of the composite model are the product of the corresponding individual functions, that is,

$$p_n(\theta) \ = \ \prod_{j=1}^k p_n^{(j)}(\theta^{(j)}) \ , \quad r(\theta) \ = \ \prod_{j=1}^k r^{(j)}(\theta^{(j)}) \ ,$$

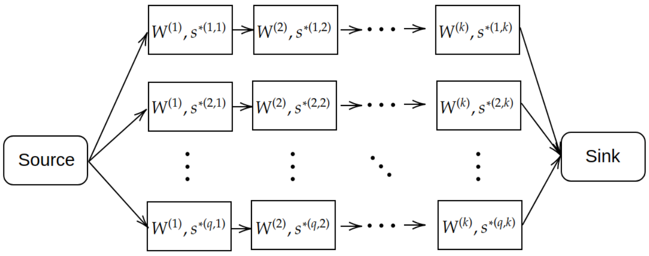

Figure 2: Diagrammatic representation of parallel hypotheses in serial models composition of a composite hypothesis, H, whose e-value, $\mathrm{ev}(H)$, can be computed from the truth-functions, $W^{(j)}$, of each single model, $M^{(j)}$, and the optimal surprise values, $s^{*(i,j)}$, of each single hypothesis, $H^{(i,j)}$.

$$s(\theta) \;=\; \prod_{j=1}^{k} s^{(j)}(\theta^{(j)}) \;, \;\; \theta = [\theta^{(j)}, \ldots \theta^{(k)}] \in \Theta \;.$$

The Mellin convolution of the cumulative distributions for (scalar) random variables x and y, $F(x) \otimes G(y)$, gives the cumulative distribution of the product $z = xy$. Hence, the composite model's truth function is the *Mellin convolution* of the individual truth functions, that is,

$$W(v) = \bigotimes_{1 \le j \le k} W^{(j)} \;=\; W^{(1)} \otimes W^{(2)} \ldots \otimes W^{(k)}(v) \;.$$

Further properties and interpretations of the Mellin convolution, and detailed analytical procedures and numerical algorithms for its efficient computation, can be found in [124, 175, 220, 233]. Matlab code for computing the composite model's truth function in the discrete case, i.e. for step ladder cumulative functions, is presented in [1].

Next, let us consider a composite hypothesis H expressed in *conjunctive form*, where the conjunction symbol ($\wedge$) stands for the *and* logical operator,

$$H = \bigwedge_{j=1}^{k} H^{(1,j)} = H^{(1,1)} \wedge H^{(1,2)} \ldots \wedge H^{(1,k)} \;.$$

In the conjunctive hypothesis, the surprise function attains a maximum (or supremum) value equal to the product of individual maxima (or suprema). Therefore, the e-value of the conjunctive composite hypothesis is given by

$$\mathrm{ev}(H) \;=\; W(s^*) \;=\; \bigotimes_{1 \le j \le k} W^{(j)} \left(\prod_{j=1}^{k} s^{*(1,j)} \right) \;.$$

Finally, let us consider a composite hypothesis H expressed in *disjunctive normal form*, where the disjunction symbol ($\vee$) stands for the *or* logical operator,

$$H \;=\; \bigvee_{i=1}^{q} \bigwedge_{j=1}^{k} H^{(i,j)} = \left(\bigwedge_{j=1}^{k} H^{(1,j)} \right) \vee \ldots \vee \left(\bigwedge_{j=1}^{k} H^{(q,j)} \right) .$$

The corresponding e-value requires the conjunctive surprise to be maximized over the finite set of q disjunctive alternatives,

$$\mathrm{ev}(H) \;=\; \mathrm{ev}\left(\bigvee_{i=1}^{q} \bigwedge_{j=1}^{k} H^{(i,j)} \right) \;=$$
$$W(s^*) \;=\; \max_{i=1}^{q} \bigotimes_{1 \le j \le k} W^{(j)} \left(\prod_{j=1}^{k} s^{*(i,j)} \right) \;.$$

Interestingly, in the last expression, if all elementary hypotheses have either null or full epistemic value, that is, if, $\forall (i,j), \mathrm{ev}(H^{(i,j)}) \in \{0,1\}$, the evaluation of $\mathrm{ev}(H)$ reduces to the the evaluation in classical logic of a corresponding expression preserving the same normal structure [2]. Another interesting aspect of the last expression is how conjunction operators are translated into product operations, while disjunction operators are translated into maximization operations. This kind of translation characterizes the e-value as a *possibilistic abstract belief calculus*. For further details and comments, see [2, 25, 26, 35, 134, 135, 139, 142, 180, 223, 224].

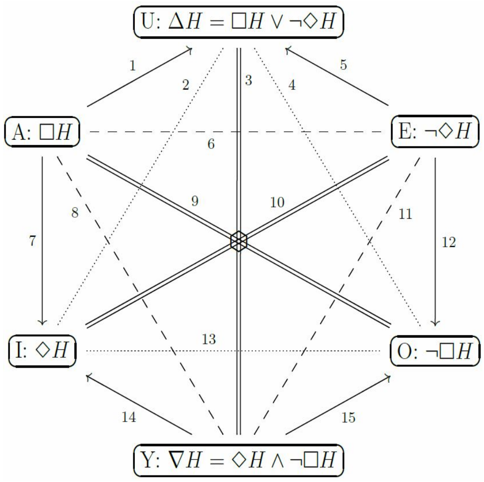

Figure 3: Hexagon of opposition depicting modal operators of necessity, $\Box$; possibility, $\Diamond$; contingency, ∇; and non-contingency, Δ; some compositions by negation, $\neg$; disjunction, $\vee$; and conjunction, $\wedge$; and logical relations of implication, $\rightarrow$; contrariety $--$; sub-contrariety, $\cdots$; and contradiction, $=$.

Name	Modality	Equivalence	Interpretation
Necessity	$\Box H$	$\Delta H \wedge \Diamond H$	H is accepted
Impossibility	$\neg\Diamond H$	$\Delta H \wedge \neg\Box H$	H is rejected
Contingency	∇H	$\Diamond H \wedge \neg\Box H$	H is not decided
Possibility	$\Diamond H$	$\Box H \vee \nabla H$	H is not rejected
Non-necessity	$\neg\Box H$	$\neg\Diamond H \vee \nabla H$	H is not accepted
Non-contingency	ΔH	$\Box H \vee \neg\Diamond H$	H is decided

Table 1:Basic modalities of agnostic hypothesis tests.

7 Statistical Hypotheses Tests and the GFBST

In the practice of statistics, it is often necessary to make a decision to either *reject*, or to remain *undecided*, or to *accept* a hypothesis H (in statistics, the alternative undecided is also called *agnostic*). In *multi-valued logics*, these three alternatives are traditionally encoded by the logical values $\{0, 1/2, 1\}$, see [147, 148, 162, 197]. In *modal logics*, these three alternative attitudes consider the hypothesis to be *impossible*, *contingent*, or *necessary*, see [126, 196, 249]. Table 1 presents the corresponding *modal operators* together with some relevant compositions obtained by negation ($\neg$, the *not* operator), disjunction ($\vee$, the *or* operator) and conjunction ($\wedge$, the *and* operator).

The modalities in Table 1 are related by logical relations depicted in diagrammatic form by the *hexagon of opposition* in Figure 3, where: Arrows ($\longrightarrow$) indicate logical *implication*; Dashed-lines ($---$) indicate *contrariety*, connecting modalities that cannot both be true; Dotted-lines ($\cdots\cdots$) indicate *sub-contrariety*, connecting modalities that cannot both be false; and Double-lines ($=$): indicate *contradiction*, connecting modalities that can neither both be true nor both be false. The vowels at the vertices of the hexagon are historical labels inherited from medieval logic, see [243].

The logical relations depicted in the hexagon of opposition are considered basic principles of rational argumentation, intuitive to, and tacitly assumed by most users and, sometimes, even soft coded in popular culture or hard coded in natural language, see [8, 43, 44, 118, 121, 243, 248]. Therefore, departing from these principles may make arguments counterintuitive, create barriers to the natural flow of communication, or even lead to misunderstandings. This situation justifies the development of inference methods and supports ways of reasoning that are fully compliant with these basic principles.

A *statistical test of hypotheses* is a statistical procedure used to make a required decision to either reject, or remain undecided, or accept a hypothesis H. It is convenient to use a statistical test based on an already available significance measure, $\mu(H) \in [0,1]$. Such a statistical test can be regarded as a discretization map, $\delta(\mu(H))$, that collapses the interval $[0,1]$ into the ternary set $\{0,1/2,1\}$. Some applications may only allow decisions in a binary subset, like $\{0,1/2\}$, although a full logical analysis of these tests requires (at least) a ternary decision space, see [6, 8, 9, 23, 41].

The remainder of this section examines logical consistency conditions for such tests, and provides further justification for the use of the e-value as an appropriate significance measure. The preceding sections paid special attention to sharp hypotheses. In contrast, this section looks at sharp or slack hypotheses with equal interest. Although motivated by the need to evaluate sharp hypotheses, we should remark that nothing prevents the e-value from being computed for slack hypotheses.

As already examined in Section 2, by definition, a statistical hypothesis H states that the true parameter value belongs to the hypothesis set, that is, that $\theta^0 \in H$. From this definition, in conjunction with classical principles or rational argumentation encoded in the hexagon of opposition, we now present three sets of conditions for logical consistency that ought to be required from a rational hypothesis test, namely, *invertibility*, *monotonicity*, and *consonance*, see [4, 5, 8, 41].

Invertibility: Every hypothesis $H \subset \Theta$ automatically defines its complement, $\overline{H} = \Theta - H$. Hence, the true value of the parameter, $\theta^0 \in \Theta$, must either belong to H or belong to its complement. Therefore, the following implications ought to hold:

(I.i) *Necessity inversion:* H is necessary if and only if $\overline{H}$ is impossible, that is, $\Box H \Leftrightarrow \neg\Diamond\overline{H}$;

(I.ii) *Possibility inversion:* H is possible if and only if $\overline{H}$ is unnecessary, that is, $\Diamond H \Leftrightarrow \neg\Box\overline{H}$; and

(I.iii) *Contingency inversion:* H is contingent if and only if $\overline{H}$ is contingent, that is, $\nabla H \Leftrightarrow \nabla\overline{H}$.

Figure 4 (left) gives a diagrammatic representation of invertibility relations (and their complements obtained by negation) using fragments of the hexagon of opposition.

Monotonicity: Let us consider a larger hypothesis H' that contains the original hypothesis H, that is $H' \supset H$. In this setting, the following implications ought to hold:

(M.i) *Monotonic necessity:* If H is necessary, so must be H', that is, $\Box H \Rightarrow \Box H'$; and

(M.ii) *Monotonic possibility:* If H is possible, so must be H', that is, $\Diamond H \Rightarrow \Diamond H'$.

Figure 4 (right) gives a diagrammatic representation of these monotonicity relations (and their complements obtained by negation) using fragments of the hexagon of opposition.

Consonance: Let H_i, for $i \in I$, be a collection of hypotheses referenced by the index set I. The following implications ought to hold:

(C.i) *Union consonance*: If the *join* hypothesis, made by the union of all indexed hypotheses, is possible, then at least one of the joining hypotheses is possible, that is,

$$\Diamond(\cup_{i\in I} H_i) \Rightarrow \exists i \in I \,|\, \Diamond H_i \ .$$

(C.ii) *Intersection consonance*: If the *meet* hypothesis, made by the intersection of all indexed hypotheses, is unnecessary, then at least one of the meeting hypotheses is unnecessary, that is,

$$\neg\Box(\cap_{i\in I} H_i) \Rightarrow \exists i \in I \,|\, \neg\Box H_i \quad \Leftrightarrow \quad \forall i \in I,\ \Box H_i \Rightarrow \Box(\cap_{i\in I} H_i) \ .$$

Figure 5 gives a diagrammatic representation of consonance relations for a collection of three hypotheses, $\{A,B,C\}$.

Arguments of analysis lead to a characterization of test procedures that comply with the aforementioned logical conditions of monotonicity, invertibility, and consonance, see [4, 9]. These tests must be based on a *region estimator*, $S \subset \Theta$, for the location of the true parameter value, θ^0. Such a *region estimator test* accepts H if $S \subset H$, rejects H if $S \subset \overline{H}$, and remains undecided if S properly intersects both H and $\overline{H}$, as depicted in Figure 6.

As mentioned in the beginning of this section, it is convenient to use a statistical test, $\delta(\mu(H)) \in \{0,1/2,1\}$, that is based on an already available significance measure, $\mu(H) \in [0,1]$. Nevertheless, the

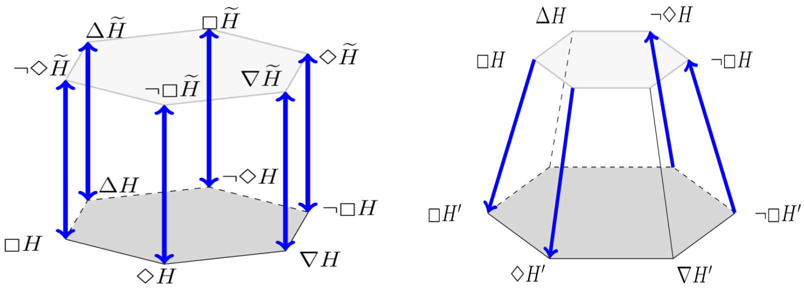

Figure 4: Invertibility (left) and Monotonicity (right) logical relations between statistical modalities applied to a hypothesis, H, and its complement, $\overline{H}$.

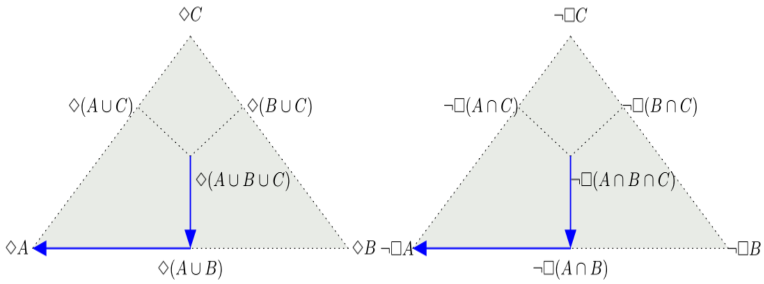

Figure 5: Consonance logical relations for the union or intersection of three hypotheses.

following logical compatibility condition, between the statistical test and its underlying significance measure, ought to hold:

$$\mu(H^{(1)}) \geq \mu(H^{(2)}) \Rightarrow \delta(\mu(H^{(1)})) \geq \delta(\mu(H^{(2)})) \; .$$

This logical compatibility condition is inspired by several legal principles, like the ones known in the juridical literature as *onus probandi*, *in dubio pro reo*, and *principle of proportionality*, see [25, 163, 201, 234, 237, 244, 250, 252, 254, 260]. The *onus probandi* and *in dubio pro reo* principles can be interpreted as requiring the monotonicity and consonance properties examined in this section, while the *principle of proportionality* can be interpreted as follows: – Judging the acceptability of hypotheses $H^{(1)}$ and $H^{(2)}$, if the former has a better grade (a higher significance measure) than the latter, then fair judgments must render verdicts (decisions) that are as least as favorable to the former than to the latter.

From the definition of the e-value, it is clear that an upper v-cut $\overline{T}(v)$ can be used as a region estimator for θ^0. Accordingly, the corresponding region estimator test rejects H if its e-value stays below the threshold $c = W(v)$, that is, H is rejected if $\text{ev}(H) < c$. By invertibility, H is accepted if $\text{ev}(\overline{H}) < c$. This is the GFBST - the *Generalized Full Bayesian Significance Test*, see [4, 7, 8, 14, 23, 24, 41]. One can check that the GFBST generates logical modalities that are compatible with all logical requirements examined in this section, and also that it is logically compatible with its underlying significance measure, the e-value. Moreover, one can also check that all good invariance and asymptotic properties of the e-value are inherited by the GFBST. Furthermore, it is interesting to remark that, under usual regularity conditions (continuity and differentiability), if $\text{ev}(H) < 1$, a sharp hypothesis H may either remain undecided or be rejected by the GFBST, but never be accepted. Finally, for simplicity and limitations of space, the exposition given in this article first defines the e-value and the GFBST, and then derives or explains their logical properties. In contrast, in [4, 5, 8, 9, 23, 41], the authors show that it is possible to go the other way around, demonstrating that the aforementioned logical properties render a complete characterization of the GFBST as a region estimator test.

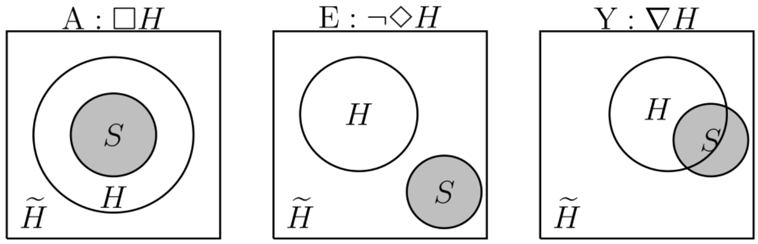

Figure 6: Testing H with region estimator S: Accept if $S \subset H$, reject if $S \subset \overline{H}$, undecided otherwise.

8 Computational Implementation

Numerical computation of the e-value can be accomplished in two basic steps, namely,

Integration step – computing an approximation of the truth function $W(v)$. In most cases, this task can be accomplished by using a computational condensation procedure, like [175, 233], on a numerical sequence, $s^{(k)} = s(\theta^{(k)})$, obtained by sampling from the posterior distribution, $p_n(\theta)$. Sampling sequences can, in turn, be generated using Markov Chain Monte-Carlo methods, see [125, 161, 166], or variations thereof, like the Hit-and-Run algorithm, see [129, 178];

Optimization step – finding the optimum of the objective function $s(\theta)$ under the constraints imposed by the hypothesis. In most cases, this task can be accomplished by standard methods of continuous constrained optimization, like Generalized Reduced Gradient, Sequential Quadratic Programming, Generalized Augmented Lagrangian, or Proximal Point methods, see [119, 151, 177, 187, 188]. The following numerical optimization solvers are readily available and potentially applicable:
TANGO – Trustable Algorithms for Nonlinear General Optimization, a joint project from USP – the Universtity of São Paulo, and UNICAMP – the University of Campinas, that offers excellent code under General Public License (GNU); and
Gurobi – a robust, high-performance, and versatile optimization software that can currently be licensed for free academic use.
Some optimization problems with many local maxima can be solved indirectly by repeated local optimization from candidate starting points selected at the integration step, or directly by stochastic optimization methods based on Simulated Annealing, see [194, 218, 221, 225, 226].

9 Numerical Example in Model Selection

Figure 7 (left) depicts the polynomial fitting problem for the classical Sakamoto et al. benchmark dataset presented in [207, ch.8]. This dataset, given in Table 1, was generated by a simulation from the i.i.d. (independent and identically distributed) stochastic process

$$y_i = g(x_i) + N(0\,,0.1^2)\ ,\ \ g(x) = \exp((x-0.3)^2) - 1\ ,$$

that is, the points y_i were generated at the grid $x_i = (i-1)0.05$, for $i = 1, \ldots 21$, by adding to the exponential *target function*, $g(x)$, a Gaussian random noise with mean $\mu = 0$ and standard deviation of $\sigma = 0.1$.

The *linear regression* polynomial model of order k explains vector y in the dataset, using as explanatory variables integer powers up to order k of the grid points in vector x, plus an i.i.d. Gaussian random noise with standard deviation σ, that is,

$$y = \beta_0 x^0 + \beta_1 x^1 + \beta_2 x^2 \ldots + \beta_k x^k + N(0, \sigma I)\ .$$

Using the weakly informative prior density, $p_0(\beta, \sigma) = 1/\sigma$, the posterior density for this statistical model can be conveniently written, see [136, 168, 235], in terms of the matrix of explanatory variables,

$X = [x^0; x^1; \ldots x^k]$, the maximum a posteriori estimators of the model's parameters, $\hat{\beta} = (X^t X)^{-1} X^t y$, and the auxiliary quantities $\hat{y} = X\hat{\beta}$ and $s^2 = (y-\hat{y})^t (y-\hat{y})/(n-k)$,

$$p_n(\beta, \sigma \,|\, y, x^0, \ldots x^k) = \frac{1}{\sigma^{n+1}} \exp\left(\frac{-1}{2\sigma^2} \left((n-k)s^2 + (\beta - \hat{\beta})^t X^t X (\beta - \hat{\beta}) \right) \right) .$$

The model selection problem that naturally arises in this context can be stated as the question: From a statistical point of view, which order k gives the best polynomial fit for this dataset?

The first column of Table 3 presents the quadratic norm *empirical error*, $R_{emp} = ||\hat{y} - y||_2$, of each model, see [136, 168, 235] for further details. Polynomials of increasing order have increasing flexibility, making R_{emp} monotonically decrease until it reaches zero for order $k = n-1$, when the model becomes just an interpolating polynomial for the dataset.

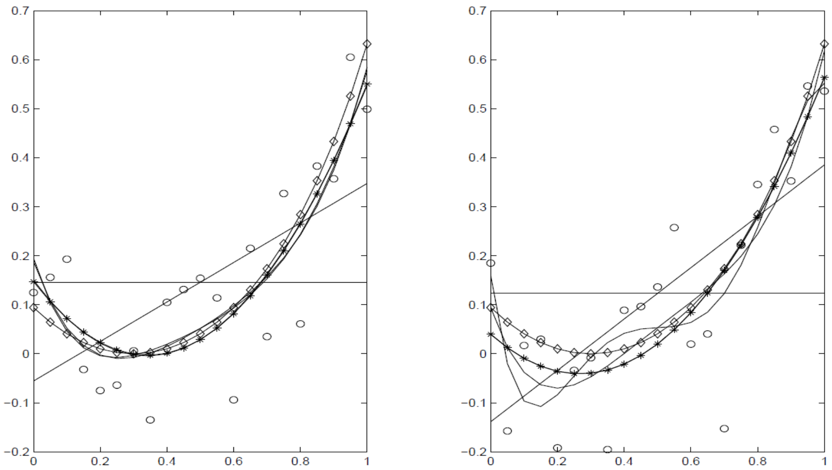

Figure 7: Benchmark (left) and alternative (right) data points, $\circ$, for Sakamoto et al. [207, ch.8] polynomial fitting problem; Exponential target function, $\diamond$; Best fitted polynomials, — , of order 0 to 4; and Order 2 polynomial, $*$, that renders the smallest regularized error for both datasets.

i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
1	0.00	0.125	8	0.35	-0.135	15	0.70	0.035
2	0.05	0.156	9	0.40	0.105	16	0.75	0.327
3	0.10	0.193	10	0.45	0.131	17	0.80	0.061
4	0.15	-0.032	11	0.50	0.154	18	0.85	0.383
5	0.20	-0.075	12	0.55	0.114	19	0.90	0.357
6	0.25	-0.064	13	0.60	-0.094	20	0.95	0.605
7	0.30	0.006	14	0.65	0.215	21	1.00	0.499

Order	R_{EMP}	R_{FPE}	R_{SBC}	R_{GCV}	R_{SMS}	R_{AIC}	ev(H)	sev(H)
0	0.03712	0.04494	0.04307	0.04535	0.04419	-07.25	0.000	0.000
1	0.02223	0.02964	0.02787	0.03025	0.02858	-20.35	0.009	0.001
2	0.01130	0.01661	0.01534	0.01724	0.01560	-32.13	0.013	0.000
3	0.01129	0.01835	0.01667	0.01946	0.01667	-30.80	0.999	0.818
4	0.01088	0.01959	0.01751	0.02133	0.01710	-29.79	0.995	0.403
5	0.01087	0.02173	0.01913	0.02445	0.01811	-27.86	0.999	0.641

Table 2: Sakamoto et al. [207, ch.8] benchmark dataset for the polynomial fitting problem

Table 3: Model selection by regularized empirical errors and ev$(\beta_k = 0 \,|\, y, [x^0; x^1; \ldots x^k])$

Nevertheless, the predictive power of these models does not monotonically increase with the decrease of empirical error. From Figure 7 (right), it is easy to see that, if the order of the polynomial is too high, the model becomes excessively complex and over-fits the data, resulting in poor generalizations (predictions) outside the grid points.

Columns 2 to 6 in Table 3 present a variety of penalized errors, $R_{pen} = r(d,n)R_{emp}$, where the regularization factor, $r(d,n)$, penalizes the statistical model complexity. Specifically, the factor, $r(d,n)$, increases with the dimension of the model's parametric space, d, relative to the number of available data points, n. For the linear regressions at hand, having parameters σ and $\beta_0, \beta_1, \ldots \beta_k$, $d = k+2$. The penalized or regularized errors in Table 3 are defined by the following regularization factors, based on the quotient $q = (d/n)$. The corresponding *model selection criteria* chose the model with smallest regularized error; see [130, 207] for theoretical or heuristic justifications and further details.

- Akaike's final prediction error, $r_{FPE} = (1+q)/(1-q)$;
- Schartz' Bayesian criterion, $r_{SBC} = 1 + \ln(n)q/(2-2q)$;
- Generalized cross validation, $r_{GCV} = (1-q)^{-2}$;
- Shibata model selector $r_{SMS} = 1+2q$.

The last columns in Table 3 present, in the context of a linear regression polynomial model of order k, the e-value and its standarized version for the statistical hypotheses stating that the parameter of higher order is null, that is, $\mathrm{ev}(H) = \mathrm{ev}(\beta_k = 0 \mid y, [x^0; x^1; \ldots x^n])$ and $\mathrm{sev}(H)$. This example demonstrates that the e-value and the FBST can replace traditional model selection criteria based on penalized prediction errors. Moreover, using a selection criterion based on the FBST renders a decision process that is invariant, asymptotically consistent, and logically coherent, as explained in the previous sections. In contrast, selection criteria based on penalized prediction errors lack even the most basic invariance properties, even though alternative model selection criteria exist that are invariant and have other desirable statistical properties, see for example [170]. Furthermore, the use of the e-value and the FBST for model selection can be extended to a variety of non-nested (separate) or nested families of hypotheses, including Bayesian classifiers, as analyzed in [12, 13, 55, 57, 69, 79, 80, 87].

10 Applications

The development of statistical significance measures and tests may be motivated by their intended theoretical properties, which, in turn, may be inspired by epistemological desiderata. Nevertheless, these significance measures and tests must also prove themselves on the battlefields of science and technology as effective, efficient, robust, and reliable tools for the trade. A collection of over a hundred published applications of the e-value has been compiled in the 2020 survey [19]. This section gives a selection of references from this survey, organized by application area.

- Testing covariance structures in multivariate Normal models, treating in a unified way several alternative hypotheses (often treated as special cases in the literature): [79, 106];
- Testing unit root and cointegration hypotheses in time series, using plain and simple forms of prior information like flat or Jeffreys priors (no need for artificial priors): [61, 63, 64, 107];
- Solving Bayesian classification problems and testing nested and non-nested or separate hypotheses: [12, 13, 51, 52, 80, 193];
- Analyzing systems' reliability from failure datasets: [73, 87, 98]
- Testing dependence structures using statistical copulas: [68];
- Testing (non)-informative sampling conditions in statistical surveys: [102];
- Model selection for generalized Poisson distributions: [69, 104];
- Model selection for generalized jump diffusion and Brownian motions, extremal distributions, and persistent memory processes: [49, 54, 78, 95, 96];
- Testing independence structures in contingency tables and multinomial models: [50, 55, 90, 192];
- Software certification according to compliance conditions: [93];
- Testing market equilibrium conditions for fundamental and financial derivative asset prices: [58];
- Testing hypotheses in empirical economic studies: [60];
- Event identification in acoustic signal processing: [70–72];

- Testing Hardy-Weinberg equilibrium in genetics: [56, 74, 81, 83, 86, 89, 97, 108];
- Testing hypotheses in biological sciences, including cases in ecology, environmental sciences, medical diagnostics, efficacy evaluation of medical procedures, psychology and psychiatry: [48, 57, 76, 82, 85, 99, 100, 103, 191];
- Testing hypotheses in astronomy and astrophysics: [59, 75].

11 Epistemology, Ontology, and Metaphysics

Historically, the development of statistical significance measures (or logical truth-values), and the corresponding tests of hypotheses used in statistical science, have been reciprocally influenced by the epistemological frameworks in which they are presented. Rev. Thomas Bayes (1701-1761), whose work was communicated posthumously by Rev. Richard Price (1723-1791), developed the first methods in this area with clear goals in mind, as stated in [238], also quoted in [236, p.84] and [40, p.245]:

> *The purpose, is to shew [show] what reason we have for believing that there are in the constitution of things, fixed laws according to which events happen, and that, therefore, the frame of the world must be the effect of the wisdom and power of an intelligent cause; and thus to confirm the argument taken from final causes for the existence of the Deity [...] It will be easy to see that the problem solved in this essay is more directly applicable to this purpose; for it shews [shows] us, with distinctness and precision, in every case of any particular order or recurrency of events, what reason there is to think that such recurrency or order is derived from stable causes or regulations in nature, and not from any of the irregularities of chance.*

The next two generations in the development of probability and statistics, led by, among others, Pierre-Simon de Laplace (1749-1827) and George Boole (1815-1864), kept these core goals unchanged. While theological questions lost interest over time, the emphasis of statistical research remained essentially metaphysical – in the (gnoseological) sense of searching for and justifying causal explanations for manifested phenomena. Moreover, these causal links were ideally expressed as natural laws in the form of precise or exact quantitative relations, see [31, 32, 40]. Karl Pearson (1857-1936) was the founder and leader of *frequentist statistics*, the dominant school of thought in statistical science during the XX century. Influenced by the *Inverted Spinozism* philosophy of Johann Gottlieb Fichte (1762-1814) and the *Positivist* ideas of Auguste Comte (1798-1857), K.Pearson radically changed the goals of statistical science. He deprecated any form of causal reasoning, the conception or verification of natural laws, or the use of metaphysical (non-observable) entities. Instead, frequentist statistics, following a strict Positivist agenda, only aims to produce good-fitted empirical models able to describe or predict directly observed quantities. Accordingly, the very use of probability calculus is restricted to variables in the sample space, and strictly forbidden for (latent) variables in the parameter space, see [40].

Bruno de Finetti (1906–1985) is responsible for the resurgence of Bayesian statistics in the second half of the XX century, see [153, 245]. De Finetti reintroduced probability calculus in the parameter space. Latter on, this expanded use of probability language proved to be very useful for it could accommodate new means and methods provided by computer science, like Markov Chain Monte Carlo and other probabilistic algorithms, see [161, 166, 198]. Nevertheless, philosophically, the de Finettian revolution was a very conservative one, remaining always amenable to the Positivist agenda. This was accomplished by making probabilities in the parameter space quantities of subjective and ephemeral character (integration variables), entities of low ontological status used only at intermediate steps in the computation of predictive probabilities for observable variables in the sample space, see [153, 245].

The *Objective cognitive constructivism* epistemological framework was developed to host the e-value, the FBST, and their formal properties, including their ability to evaluate and test sharp statistical hypotheses. It also provides a naturalized way to ontology and metaphysics in empirical sciences via statistics, that is, a natural way to evaluate empirical support for natural laws and their accompanying causal explanations, and to validate the "objective" use of non-directly observable (or metaphysical) entities, in accordance with the goals of the original works of Bayes and Laplace. For further details on the Objective cognitive constructivism epistemological framework and the way in which, on the one hand, it accommodates the e-value, the FBST and their logical properties, and, on the other hand, it provides a naturalized approach to ontology and metaphysics, see [27–29, 31–33, 36–38, 40–44].

12 Future Research and Final Remarks

In December 2018, at CLE-UNICAMP, Walter Carnielli organized the workshop *Induction, Probability and their Dilemmas*, where the first author gave the presentation *The Problem of Induction in Statistical Science*. During this workshop, we discussed several topics for further research, most of them motivated by areas of common interest that, nevertheless, are usually approached quite differently by the communities of Logic and Statistics. Four of these topics are presented in the sequel, followed by some additional topics for further research.

Statistics in (Un)Countable Sentential Probability

There is a long standing tradition in Logic to formalize probabilistic reasoning over finitary or countable sets of sentences in a language. This approach is appealing for its theoretical simplicity, for allowing computer-efficient implementations of specific models, etc. For historical analyzes of this approach, see [133, 164, 165, 205], for recent works from Walter's group following this line, see [200, 239]. Notwithstanding their usefulness, simplicity, and popularity, finitary or countable sentential formalisms also have their limitations, for example, being unable to express measure-theoretic arguments used in mathematical statistics.

Sections 6 and 7 sowed how the statistical definitions of the e-value and the GFBST characterize their logical properties, and hinted at how arguments of mathematical analysis enable one to travel the other way around. In this context, it is a topic for further research to explore how much of this theory and its applications can be expressed using the sentential probability approach, either by expanding the underlying languages (like infinitary, second order, or fixed-point logics), see [146, 158, 159, 164, 165, 167, 169], or by exploring relevant properties in specific statistical models (including topological properties, like countability and compactness, and regularity conditions of constraints and distributions, like bounded continuity and differentiability), see [2, 11, 144, 153, 186, 210].

Functional Compositionality Structures

As already noticed in Section 6, the compositional structure of e-values can be characterized as a *possibilistic abstract belief calculus*, as analyzed in [2, 2, 25, 35, 134, 135, 139, 142, 180]. This particular algebraic formalism resembles in many ways the formalism used in statistical reliability theory for the analysis of complex systems assembled by serial/ parallel composition of simple(r) elements, see [114, 120, 179, 214]. As noticed in [2, Sec.1], this formal resemblance was a source of inspiration at early stages of this research program. At the same time, the truth-function interpretation of the cumulative surprise function, $W(v)$, was inspired by the work of Ludwig Wittgenstein [242, 264]. Finally, the concept of pragmatic hypotheses and its use in conjunction with the GFBST in testing scientific theories was inspired by sensitivity analysis, as used in optimization and systems' theory, and its interpretation in terms of fuzzy or paraconsistent logics, see [5, 26]. While the many theoretical results already obtained in this research program justify (in our opinion) the analogies we made and the conceptual links we established between fragments of formal structures used in distinct (and often faraway) research areas, we believe that these intuitive connections deserve and can benefit from an even more rigorous and general setting. In this context, the tools of category theory or other formal abstraction methods, see [126, 127, 155, 183, 219], offer an opportunity for further research in the study of basic logical properties of the aforementioned systems, specially concerning the investigation of their essential compositionality structures and possible generalizations.

Rough and Fuzzy Sets

Essential logical properties of the GFBST can be explained regarding the e-value as a transformation between probability and possibility measures, see [35]. The underlying interconnections between alternative representations of uncertainty at the core of this theory provide a general motivation to more specific topics of further research presented in this section.

In the GFBST framework, a sharp hypothesis can be either rejected or remain undecided, but can never be accepted, see Section 7. Aiming to allow the acceptance of surrogate hypotheses of interest, the authors showed how to enlarge an underlying sharp hypothesis into a slack *pragmatic hypothesis*, see [5]. This is accomplished by taking into account (im)precisions of measurement equipment, (un)certainties of fundamental constants, and other relevant metrological or methodological error bounds. It should be remarked that the pragmatic hypotheses defined in [5] are crisp sets.

In the context of exact sciences, the sharp (and, hence, crisp) nature of the original statistical hypothesis of interest is well-supported by the *Objective cognitive constructivism* epistemological framework; see [27–29, 31–33, 36–38, 40–42]. In contrast, the definition of the associated pragmatic hypothesis as a crisp set may be considered an over-simplification. Rough and fuzzy sets offer a theoretical framework that allows the construction of generalized pragmatic hypotheses, see [139–142, 180, 181, 208]. Rough or fuzzy pragmatic hypotheses should be able to overcome artificial limitations imposed by crisp set representations, support natural and intuitive interpretations in varied contexts of application, and still preserve the best logical properties of the GFBST.

The theoretical framework of paraconsistent logic can be used to interpret sensitivity analyses of the e-value concerning changes in the statistical model's prior or reference measures, see [26]. The same article shows how to integrate such sensitivity tests into (crisp) confidence intervals. Following the same rationale used in the last paragraph, these crisp intervals could be generalized to rough or fuzzy sets able to provide better representations of pertinent uncertainties.

Law, Complexity, and (In)Consequence in Social Systems

Early scientists used the expression *natural law* as a metaphor, that ferried to the domain of nature the normative character that a law has over human behavior. Now that mankind learned a good deal about natural laws, including their mathematical, logical, epistemic, ontological, and metaphysical characteristics, we could perhaps travel the metaphorical path taken by early scientists in the opposite direction. We could do so in an attempt to use lessons learned about natural laws to better understand the nature of laws intended to establish norms for human behavior, relations, and interactions in social systems. We believe this task can be accomplished within the framework provided by Niklas Luhmann (1927–1998) *Sociological theory of law*, see [39, 240, 246, 247, 255–257, 261, 263, 265].

In Luhmann's theory of law, the main purpose of the legal system is *congruent generalization of normative behavior expectations*, see [255, pp.77,82]. This emblematic statement can be interpreted as follows, see [39]: In Luhmann's view, norms are neither preexisting conditions nor a priori factual realities. Instead, norms are conceived as intentional projects or idealized models of how society should be; see [255, p.40]. Moreover, in such idealized models, social harmony is based on establishing well-defined, stable, sustainable, and reliable behavioral patterns, also known as *eigen-behaviors* or *eigen-solutions*, see [39, 247]. Furthermore, laws and regulations of a society should reflect its norms, stimulating/ penalizing forms of conduct that sustain/ disrupt virtuous eigen-behaviors. Finally, in Luhmann's view, social norms are not static but essentially dynamic, co-evolving (over long runs) with the behavioral patterns in the society that they simultaneously try to describe, regulate and stabilize (in their present forms, over short runs). In this context, we can try to reinterpret some lessons about model selection and their complexity learned in Section 9, claiming that, as in the case of good empirical laws, good social laws should follow the golden path of equilibrium, avoiding extremes of scarcity and excess, see also [28, 32].

On the one hand, oversimplified social laws fail to capture important distinctions considered necessary or relevant for establishing sustainable eigen-behaviors. On the other hand, excessively complex legislation creates all sorts of misinterpretations, unforeseen loopholes, and other unintended consequences. Moreover, such spurious side effects may not only obstruct virtuous eigen-behaviors, but even induce vicious ones (that must then be detected, identified, and inhibited). Furthermore, increasing legal complexities imply increasing processing times, delayed justice, and greater economic costs to operate the legal system – all burdens to be paid by the society the same system serves.

The areas of Logic and Computer science have developed several methods to measure computational complexity, either by counting processing operations in an algorithm or by accessing the code length of its description, see [146, 149, 172, 199, 227, 228]. Some of the methods used to measure complexity in statistical models have already been explained in Section 9. Meanwhile, Information science and Systems' theory have focused on entropy related measures of complexity, see [31, 34, 105, 109, 113, 143, 176]. In contrast, the mathematical treatment of complexity in theoretical and empirical legal studies is still incipient, see [203, 204, 215, 215, 241, 253, 258, 259, 262]. Naïve adaptations of complexity measures artificially borrowed from other areas often fail to capture relevant aspects of the legal environment.

Although this topic of further research can potentially benefit from all the sophisticated technical developments and approaches to complexity theory already mentioned, it also requires a healthy dose of critical thinking applied to the human condition in daily life, yet another area of interest of Walter's research group, see [128].

Individual vs. Collective Liability, Legal Burden of Proof, etc.

As commented at the end of Section 2, the *Onus Probandi* and *In Dubio Pro Reo* principles, as they are used in the legal system, were a source of inspiration for the definitions of the e-value and the FBST, see [25, 39, 93, 163, 190, 250, 254, 260]. Moreover, it has long been recognized that many decision procedures traditionally used in Bayesian statistics are incompatible with these principles. Dennis Lindley uses the famous *gatecrasher* example to argue that, under certain circumstances, collective responsibility is a preferable principle of justice, and that legal decision procedures implying collective liabilities and punishments should be accepted, see [132, 185]. Furthermore, Lindley shows how legal decisions based on Bayes Factors are compatible with such a collective liability principle.

The right tool for the right job! In accordance with this dictum, further research could contrast the legal and social implications of alternative liability principles, their potential or desirability, and their compatibility with alternative inference theories, decision procedures, and epistemological frameworks.

e-value and FBST under Non-Standard Regularity Conditions

The asymptotic consistency and standardization procedures presented in Section 5 rely on standard regularity conditions that include continuity and differentiability of several functions defining the statistical model, and also the location of the true value of the parameter at the topological interior of the parameter space. However, many interesting models present non-standard conditions. For example, mixture models and separate hypotheses models often present hypotheses at the (topological) border of the parameter space, see [12, 13], whereas eigenvalue and eigenvector inference problems may present far more challenging conditions, see [209]. The methods presented in [131, 138, 211] and references therein provide tools to generalize the e-value standardization procedure and the corresponding asymptotic results to non-standard regularity conditions.

Other non-standard inference conditions arise in semi-parametric models like, for example, Hilbert space models based on Fourier, orthogonal polynomials, wavelets, and other infinite functional bases. In these cases, it is often necessary to use weakly informative priors that dampen high-frequency modes that effectively reduce the size of resulting truncated models, see [42, Sec.6] and references therein. Further research should explore convenient forms and characteristics of such weakly informative priors and related topics.

Efficient Computational Methods and Implementations

The e-value and FBST research program would greatly benefit from computational tools tailor-made for computational efficiency, reliability, and ease of use in the numerical tasks described at Section 8. Most beneficial would be the construction of an open-source and user-friendly environment combining specially adapted versions of the following techniques: (1) Efficient Monte-Carlo integrators based on Hit-and-Run, Nested Sampling and similar algorithms, see [129, 178, 216, 217]; (2) Efficient computational condensation procedures used in conjunction with the aforementioned Monte-Carlo procedures, see [175, 233]; (3) Efficient stochastic optimization methods working in conjunction with the aforementioned Monte-Carlo procedures, see [194, 218, 221, 225, 226]; (4) Higher order asymptotic approximations used either to develop stand-alone fast e-value calculation procedures or to develop variance reduction techniques for the aforementioned Monte-Carlo procedures, see [3, 20–22, 45–47, 94]; (5) Efficient numerical convolution procedures, used in conjunction with the aforementioned condensation procedures, as required in the analysis of complex models as described in Section 6.

Acknowledgments

The authors take this opportunity to thank Prof. Walter Carnielli for, along the years, inviting, welcoming, and stimulating investigations, discussions, and interactions at areas of overlap between logic, probability, statistics, formal methods in science, philosophy of science, and general philosophy. The authors have benefited greatly from Walter Carnielli's encyclopedic knowledge, open-mindedness, positive attitude, and active engagement in fostering research in all these areas, either at his home base at CLE - the Center for Logic, Epistemology, and History of Science of UNICAMP, or throughout Brazil and around the world. The authors are grateful to Profs. Abilio Rodrigues, Alfredo Freire and Henrique Antunes for organizing a Festschrift celebrating Walter's 70th birthday.

The authors thank the organizers of the following conferences, where this paper was presented: CLMPST-2023, the 17th Congress on Logic, Methodology and Philosophy of Science and Technology, at FCE-UBA, Buenos Aires, July 24-29, 2023; Principia-2023, the 13th Principia International Symposium, at UFSC, Florianópolis, August 14-17, 2023; RatioLog-2024, the 2nd Workshop on Logic, Rationality, and Probability, at CLE-Unicamp, Campinas, April 22-25, 2024; and MaxEnt-2024, the 43rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, at U-Ghent, July 1-5, 2024.

Support, Funding, and Conflicts of Interest

The authors have received direct or indirect support from the following institutions: USP – University of São Paulo (JMS, CABP, MSL, LGE, and WSB); UFSCar – Federal Univesity of São Carlos (RI, RBS); ABJ – Brazilian Jurimetrics Association (JMS, RBS); CNPq – Brazilian National Counsel of Technological and Scientific Development, grants PQ 303290/2021-8 (JMS), PQ 302767/2017-7 (CABP), and PQ 309607-2020-5 (RI); FAPESP – State of São Paulo Research Foundation, grants 2019/11321-9 (RI), CEPID CeMEAI – Center for Mathematical Sciences Applied to Industry 2013/07375-0 (JMS, MSL), and CEPID Shell-RCGI – Research Center for Greenhouse Gas Innovation 2014/50279-4 (JMS); and CAPES – Coordination for Improvement of Higher Education Personnel, finance code 001 (RI, RBS, MAD). Concerning this article, the authors have no conflicts of interest to declare.

References

– e-value Theory and Epistemology

[1] Borges, Wagner de Souza; Stern, Julio Michael. Evidence and Compositionality. pp.307-315 in Lawry, J. et al. *Soft Methods for Integrated Uncertainty Modelling*, Berlin: Springer. Proceedings of SMPS-2006.

[2] Borges, Wagner de Souza; Stern, Julio Michael. The Rules of Logic Composition for the Bayesian Epistemic E-Values. *Logic Journal of the IGPL*, 15, 5/6, 401-420, 2007.

[3] Cabras, Stefano; Racugno, Walter; Ventura, Laura. Higher order asymptotic computation of Bayesian significance tests for precise null hypotheses in the presence of nuisance parameters. *Journal of Statistical Computation and Simulation*, 85, 15, 2989-3001, 2015.

[4] Esteves, Luis Gustavo; Izbicki, Rafael; Stern, Julio Michael; Stern, Rafael Bassi. The logical consistency of simultaneous agnostic hypothesis tests. *Entropy*, 18, 256, 2016

[5] Esteves, Luis Gustavo; Izbicki, Rafael; Stern, Julio Michael; Stern, Rafael Bassi. Pragmatic Hypotheses in the Evolution of Science. *Entropy*, 21, 9, 883, 2019.

[6] Esteves, Luis Gustavo ; Izbicki, Rafael; Stern, Julio Michael; Stern, Rafael Bassi. Logical Coherence in Bayesian Simultaneous Three-Way Hypothesis Tests. *International Journal of Approximate Reasononig*, 152, 297-309, 2023.

[7] Fossaluza, Victor; Izbicki, Rafael; Silva, Gustavo Miranda da; Esteves, Luís Gustavo. Coherent Hypothesis Testing. *The American Statistician*, 71, 3, 242-248, 2017.

[8] Izbicki, Rafael; Esteves, Luis Gustavo. Logical consistency in simultaneous statistical test procedures. *Logic Journal of the IGPL*, 23, 732-758, 2015.

[9] Izbicki, Rafael; Stern, Julio Michael; Stern, Rafael Bassi; Esteves, Luís Gustavo. Logical Coherence in

Bayesian Simultaneous Three-Way Hypothesis Tests. Submitted for publication, 2022.

[10] Kelter, Riko; Stern, Julio Michael. The Full Bayesian Significance Test and the e-value – Foundations, theory and application in the cognitive sciences, 2020. `arXiv:2006.03334`.

[11] Kelter, Riko. On the Measure-Theoretic Premises of Bayes Factor and Full Bayesian Significance Tests: A Critical Reevaluation. *Computational Brain & Behavior*, 2021. `doi:10.1007/s42113-021-00110-5`.

[12] Lauretto, Marcelo de Souza; Stern, Julio Michael. FBST for Mixture Model Selection. *AIP Conference Proceedings*, 803, 121-128, 2005.

[13] Lauretto, Marcelo de Souza; Faria, Silvio; Pereira, Basilio de Bragança; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. The Problem of Separate Hypotheses via Mixtures Models. *AIP Conference Proceedings*, 954, 268-275, 2007.

[14] Madruga, Maria Regina; Esteves, Luis Gustavo; Wechsler, Sergio. On the Bayesianity of Pereira-Stern Tests. *Test*, 10, 291-299, 2001.

[15] Madruga, Maria Regina; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. Bayesian Evidence Test for Precise Hypotheses. *Journal of Statistical Planning and Inference*, 117, 185-198, 2003.

[16] Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses. *Entropy*, 1, 99-110, 1999.

[17] Pereira, Carlos Alberto de Bragança; Stern, Julio Michael; Wechsler, Sergio. Can a Significance Test be Genuinely Bayesian? *Bayesian Analysis*, 3, 79-100, 2008.

[18] Pereira, Carlos Alberto de Bragança. Full Bayesian Significant Test (FBST). pp.551-554 in Lovric, M. ed. *International Encyclopedia of Statistical Science*, Berlin, Springer, 2011.

[19] Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. The e-value: A fully Bayesian significance measure for precise statistical hypotheses and its research program. *São Paulo Journal of Mathematical Sciences*, 2020. `doi:10.1007/s40863-020-00171-7`.

[20] Pinto, Anna; Ventura, Laura. *Approssimazioni Asintotiche di Ordine Elevato per Verifiche d'Ipotesi Bayesiani: Uno Studio per Dati di Sobrevvivenza*. Università degli Studi di Padova, Dipartimento di Scienze Statistiche, 2012.

[21] Ruli, Erlis; Sartori, Nicola; Ventura, Laura. Robust approximate Bayesian inference. *Journal of Statistical Planning and Inference*, 205, 10-22, 2020.

[22] Ruli, Erlis; Ventura, Laura. Can Bayesian, Confidence Distribution and Frequentist Inference Agree? *Statistical Methods and Applications*, 2020. `doi:10.1007/s10260-020-00520-y`.

[23] Silva, Gustavo Miranda da; Esteves, Luis Gustav; Fossaluza, Victor; Izbicki, Rafael; Wechsler, Sergio. A Bayesian Decision Theoretic Approach to Logically-Consistent Hypothesis Testing. *Entropy*, 17, 10, 6534-6559, 2015.

[24] Silva, Ivair R. On the correspondence between frequentist and Bayesian tests. *Communications in Statistics - Theory and Methods*, 47, 14, 3477-3487, 2018.

[25] Stern, Julio Michael. Significance Tests, Belief Calculi, and Burden of Proof in Legal and Scientific Discourse. 4th Laptec – Logic Applied to Technology congress, Advances in Intelligent Systems and Robotics. *Frontiers in Artificial Intelligence and its Applications*, 101, 139-147, 2003.

[26] Stern, Julio Michael. Paraconsistent Sensitivity Analysis for Bayesian Significance Tests. *Lecture Notes in Artificial Intelligence*, 3171, 134-143, 2004.

[27] Stern, Julio Michael. Cognitive Constructivism, Eigen-Solutions, and Sharp Statistical Hypotheses. *Cybernetics & Human Knowing*, 14, 1, 9-36, 2007.

[28] Stern, Julio Michael. Language and the Self-Reference Paradox. *Cybernetics & Human Knowing*, 14, 4, 71-92, 2007.

[29] Stern, Julio Michael. Decoupling, Sparsity, Randomization, and Objective Bayesian Inference. *Cybernetics & Human Knowing*, 15, 2, 49-68, 2008.

[30] Stern, Julio Michael. *Cognitive Constructivism and the Epistemic Significance of Sharp Statistical Hypotheses in Natural Sciences*. Tutorial text for MaxEnt 2008 - The 28th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Boracéia, São Paulo, Brazil, July 6-11, 2008. `arXiv:1006.5471`.

[31] Stern, Julio Michael. Symmetry, Invariance and Ontology in Physics and Statistics. *Symmetry*, 3, 3, 611-635, 2011.

[32] Stern, Julio Michael. Constructive Verification, Empirical Induction, and Falibilist Deduction: A Threefold Contrast. *Information*, 2, 635-650, 2011.

[33] Stern, Julio Michael. Jacob's Ladder and Scientific Ontologies. *Cybernetics & Human Knowing*, 21, 3, 9-43, 2014.

[34] Stern, Julio Michael; Nakano, Fabio. Optimization Models for Reaction Networks: Information Divergence, Quadratic Programming and Kirchhoff's Laws. *Axioms*, 3, 109-118, 2014.

[35] Stern, Julio Michael; Pereira, Carlos Alberto de Bragança. Bayesian Epistemic Values: Focus on Surprise, Measure Probability! *Logic Journal of the IGPL*, 22, 236-254, 2014.

[36] Stern, Julio Michael. Cognitive-Constructivism, Quine, Dogmas of Empiricism, and Münchhausen's Trilemma. Ch.5, pp.55-68, in Polpo, Adriano; Louzada, Francisco; Rifo, Laura L.R; Stern, Julio M; Lauretto, Marcelo; eds. Interdisciplinary Bayesian Statistics: EBEB 2014, *Springer Proceedings in Mathematics and Statistics*, 118, 2015.

[37] Stern, Julio Michael. Continuous versions of Haack's Puzzles: Equilibria, Eigen-States and Ontologies. *Logic Journal of the IGPL*, 25, 4, 604-631, 2017.

[38] Stern, Julio Michael. Jacob's Ladder: Logics of Magic, Metaphor and Metaphysics – Narratives of the Unconscious, the Self, and the Assembly. *Sophia*, published at Springer Online First June 7, 2017. doi:10.1007/s11841-017-0592-y.

[39] Stern, Julio Michael. Verstehen (causal/ interpretative understanding), Erklären (law-governed description/ prediction), and Empirical Legal Studies. *Journal of Institutional and Theoretical Economics* (JITE), 174, 105-114, 2018.

[40] Stern, Julio Michael. Karl Pearson on Causes and Inverse Probabilities: Renouncing the Bride, Inverted Spinozism and Goodness-of-Fit. *South American Journal of Logic*, 4, 1, 219-252, 2018.

[41] Stern, Julio Michael; Izbicki, Rafael; Esteves, Luis Gustavo; Stern, Rafael Bassi. Logically-Consistent Hypothesis Testing and the Hexagon of Oppositions. *Logic Journal of the IGPL*, 25, 741-757, 2018.

[42] Stern, Julio Michael. A Sharper Image: The Quest of Science and Recursive Production of Objective Realities. *Principia*, 24, 2, 255-297, 2020.

[43] Stern, Julio Michael. Color-Coded Epistemic Modes in a Jungian Hexagon of Opposition. In Jean-Yves Beziau, Ioannis Vandoulakis (2022), *Studies in Universal Logic*, 20, 303-332, 2022. doi:10.1007/978-3-030-90823-2_14

[44] Stern, Julio Michael. Dynamic Oppositional Symmetries for Color, Jungian and Kantian Categories. *Logica Universalis*, 2023. doi:10.1007/s11787-023-00342-y

[45] Ventura, Laura; Ruli, Erlis; Racugno, Walter. A note on approximate Bayesian credible sets based on modified loglikelihood ratios. *Statistics and Probability Letters*, 83, 11, 2467-2472, 2013.

[46] Ventura, Laura; Reid, Nancy. Approximate Bayesian computation with modified log-likelihood ratios. *Metron*, 72, 231-245, 2014.

[47] Ventura, Laura; Racugno, Walter. Pseudo-Likelihoods for Bayesian Inference. pp.205-220 in DiBattista, T; Moreno, E; Racugno, W; eds. *Topics on Methodological and Applied Statistical Inference*. Berlin, Springer, 2016.

– e-value Use in Empirical Science

[48] Ainsbury, Elizabeth A; Vinnikov, Volody-myr A; Puig, Pedro; Higueras, Manuel; Maznyk, Nataliya A; Lloyd, David C; Rothkamm, Kai. Review of Bayesian statistical analysis methods for cytogenetic radiation biodosimetry, with a practical example. *Radiation Protection Dosimetry*, 162, 3, 185-196, 2013.

[49] Andrade, Plinio; Rifo, Laura Leticia Ramos; Torres, Soledad; Torres-Avilés, Francisco. Bayesian Inference on the Memory Parameter for Gamma-Modulated Regression Models. *Entropy*, 17, 10, 6576-6597, 2015.

[50] Andrade, Pablo De Morais; Stern, Julio Michael; Pereira, Carlos Alberto De Bragança. Bayesian Test of Significance for Conditional Independence: The Multinomial Model. *Entropy*, 16, 3, 1376-1395, 2014.

[51] Assane, Cachimo Combo; Pereira, Basilio de Bragança; Pereira, Carlos Alberto de Bragança. Bayesian significance test for discriminating between survival distributions. *Communications in Statistics - Theory and Methods*, 47, 24, 6095-6107, 2018.

[52] Assane, Cachimo Combo; Pereira, Basilio de Bragança; Pereira, Carlos Alberto de Bragança. Model choice in separate families: A comparison between the FBST and the Cox test. *Communications in Statistics - Simulation and Computation*, 48, 9, 2641-2654, 2019.

[53] Barahona, Manuel; Rifo, Laura; Sepúlveda, Maritza; Torres, Soledad. A Simulation-Based Study on Bayesian Estimators for the Skew Brownian Motion. *Entropy*, 18, 7, 241, 1-14, 2016.

[54] Bernardini, Diego F. de; Rifo, Laura Leticia Ramos. Full Bayesian significance test for extremal distributions. *Journal of Applied Statistics*, 38, 4, 851-863, 2011.

[55] Bernardo, Gustavo; Lauretto, Marcelo de Souza; Stern, Julio Michael. The Full Bayesian Significance Test for Symmetry in Contingency Tables. *AIP Conference Proceedings*, 1443, 198-205, 2012.

[56] Brentani, Helena; Nakano, Eduardo Y; Martins, Camila B; Izbicki, Rafael; Pereira, Carlos Alberto. Disequilibrium Coefficient: A Bayesian Perspective. *Statistical Applications in Genetics and Molecular Biology*, 10, 1, 22, 1-24, 2011.

[57] Camargo, André P; Stern, Julio M; Lauretto, Marcelo S. Estimation and model selection in Dirichlet regression. *AIP Conference Proceedings*, 1443, 206-213, 2012.

[58] Cerezetti, Fernando Valvano; Stern, Julio Michael. Non-arbitrage in Financial Markets: A Bayesian Approach for Verification. *AIP Conference Proceedings*, 1490, 87-96, 2012.

[59] Chakrabarty, Dalia. A New Bayesian Test to Test for the Intractability-Countering Hypothesis. *Journal of the American Statistical Association*, 112, 518, 561-577, 2017.

[60] Chaiboonsri, Chukiat; Wannapan, Satawat; Saosaovaphak, Anuphak. Economic and Business Cycle of India: Evidence from ICT Sector. p.29-43 in Tsounis, Nicholas; Vlachvei, Aspasia. *Advances in Panel Data Analysis in Applied Economic Research*. Cham, Switzerland, Springer Nature, 2018.

[61] Chen, C. W. S; Lee, S. A local unit root test in mean for financial time series. *Journal of Statistical Computation and Simulation*, 86, 4, 788-806, 2015.

[62] D'Cunha, Juliet Gratia; Rao, Aruna K. Frequentist Comparison of the Bayesian Significance Test for Testing the Median of the Lognormal Distribution. *InterStat*, 2016, 02, 001, 1-25, 2016.

[63] Diniz, Marcio; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. Unit Roots: Bayesian Significance Test. *Communications in Statistics - Theory and Methods*, 40, 23, 4200-4213, 2011.

[64] Diniz, Marcio; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. Cointegration: Bayesian Significance Test. *Communications in Statistics - Theory and Methods*, 41, 19, 3562-3574, 2012.

[65] Diniz, Marcio; Pereira, Carlos A.B; Polpo, Adriano; Stern, Julio M; Wechsler, Sergio. Relationship between Bayesian and Frequentist Significance Indices. *International Journal for Uncertainty Quantification*, 2, 2, 161-172, 2012.

[66] Diniz, Marcio; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. *e-value* entry at Wiley's *StatsRef: Statistics Reference*, 2022. `doi:10.1002/9781118445112.stat08375`

[67] Fossaluza, Victor; Esteves, Luís Gustavo; Pereira, Carlos Alberto de Bragança. Estimating Multivariate Discrete Distributions Using Bernstein Copulas. *Entropy*, 20, 3, 194, 1-16, 2018.

[68] García, Jesús E; González-López, Verónica; Nelsen, Roger B. The Structure of the Class of Maximum Tsallis-Havrda-Chavat Entropy Copulas. *Entropy*, 18, 7, 264, 1-6, 2016.

[69] Hubert, Paulo; Lauretto, Marcelo de Souza; Stern, Julio Michael 2009. FBST for Generalized Poisson Distribution. *AIP Conference Proceedings*, 1193, 210-2019, 2009.

[70] Hubert, Paulo; Padovese, Linilson; Stern, Julio Michael. A Sequential Algorithm for Signal Segmentation. *Entropy*, 20, 1, 55, 1-20, 2018.

[71] Hubert, Paulo; Stern, Julio Michael. Probabilistic Equilibrium: A Review on the Application of MaxEnt to Macroeconomic Models. *Springer Proceedings in Mathematics and Statistics*, 239, 187-197, 2018.

[72] Hubert, Paulo; Killick, Rebecca; Chung, Alexandra; Padovese, Linilson R. A Bayesian binary algorithm for root mean squared-based acoustic signal segmentation. *Journal of the Acoustical Society of America*, 146, 3, 1799-1807, 2019.

[73] Irony, Telba Zalkind; Lauretto, Marcelo de Souza; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. A Weibull Wearout Test: Full Bayesian Approach. pp.287-300 in Hayakawa, Y; Irony, T.Z; Xie, M; eds. *Systems and Bayesian Reliability*. Singapore, World Scientific, 2002.

[74] Izbicki, Rafael; Fossaluza, Victor; Hounie, Ana Gabriela; Nakano, Eduardo Yoshio; Pereira, Carlos Alberto de Bragança. Testing allele homogeneity: the problem of nested hypotheses. *BMC Genetics*, 13, 103, 1-11, 2012.

[75] Johnson, Ria; Chakrabarty, Dalia; O'Sullivan, Ewan; Raychaudhury, Somak. Comparing X-Ray and Dynamical Mass Profiles in the Early-Type Galaxy NGC 4636. *The Astrophysical Journal*, 706, 2, 706, 980-994, 2009.

[76] Kelter, Riko. Analysis of Bayesian posterior significance and effect size indices for the two-sample t-test to support reproducible medical research. *BMC Medical Research Methodology*, 20, 88, 1-18, 2020

[77] Kelter, Riko. Bayesian Hodges-Lehmann tests for statistical equivalence in the two-sample setting: Power analysis, type I error rates and equivalence boundary selection in biomedical research. *BMC Medical Research Methodology*, 21, 171, 1-26, 2021.

[78] Kostrzewski, Maciej. On the Existence of Jumps in Financial Time Series. *Acta Physica Polonica B*, 43, 10, 2001-2019, 2012.

[79] Lauretto, Marcelo de Souza; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael, Zacks, Shelemyahu. Full Bayesian Significance Test Applied to Multivariate Normal Structure Models. *Brazilian. Journal of Probability and Statistics*, 17, 147-168, 2003.

[80] Lauretto, Marcelo de Souza; Stern, Julio Michael. Testing Significance in Bayesian Classifiers. *Frontiers in Artificial Intelligence and Applications*, 132, 34-41, 2005.

[81] Lauretto, Marcelo de Souza; Nakano, Fabio; Faria, Silvio; Pereira, Carlos A. de Bragança; Stern, Julio Michael. A Straightforward Multiallelic Significance Test for the Hardy-Weinberg Equilibrium Law. *Genetics and Molecular Biology*, 32, 3, 619-625, 2009.

[82] Lima, Adriano R; Mello, Marcelo; Andreoli, Sérgio; Fossaluza, Victor; Araújo, Célia de; Jackowski, Andrea; Bressan, Rodrigo; Mari, Jair. The Impact of Healthy Parenting As a Protective Factor for Posttraumatic Stress Disorder in Adulthood: A Case-Control Study. *PLOS ONE*, 9, 1, 1-9, e87117, 2014.

[83] Loschi, Rosangela Helena; Monteiro, João V.D; Rocha, Gustavo H.M.A; Mayrink, Vinicius D. Testing and Estimating the Non-Disjunction Fraction in Meiosis I using Reference Priors. *Biometrical Journal*, 49, 6, 824-839, 2007.

[84] Loschi, Rosangela Helena; Santos, Cristiano C; Arellano-Valle, Reinaldo B. Test procedures based on combination of Bayesian evidences for H_0. *Brazilian J. of Probability and Statistics*, 26, 4, 450-473, 2012.

[85] Mathis, Maria Alice de; Rosario, Maria C.do; Diniz, Juliana Belo; Torres, Albina R; Shavitt, Roseli G; Ferrão, Ygor A; Fossaluza, Victor; Pereira, C.A.B; Miguel, Eurípedes Constantino. Obsessive-compulsive disorder: Influence of age at onset on comorbidity patterns. *European Psychiatry*, 23, 3, 187-194, 2008.

[86] Montoya-Delgado, Luis E; Irony, Telba Zalkind; Pereira, Carlos A. de B; Whittle, Martin R. An Unconditional Exact Test for the Hardy-Weinberg Equilibrium Law: Sample-Space Ordering Using the Bayes Factor. *Genetics*, 158, 2, 875-883, 2001.

[87] Maranhão, Viviane de Luca; Lauretto, Marcelo de Souza; Stern, Julio Michael. FBST for Covariance Structures of Generalized Gompertz Models. *AIP Conference Proceedings*, 1490, 202-211, 2012.

[88] Marcondes, Diego; Peixoto, Peixoto; Stern, Julio Michael. Assessing randomness in case assignment: The case study of the brazilian supreme court. *Law, Probability and Risk*, 18, 2-3, 97-114, 2019.

[89] Nakano, Fabio; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael; Whittle, Martin R. Genuine Bayesian Multiallelic Significance Test for the Hardy-Weinberg Equilibrium Law. *Genetics and Molecular Research*, 4, 619-631, 2006.

[90] Oliveira, Natalia L; Pereira, Carlos A.B; Diniz, Marcio A; Polpo Adriano. A discussion on significance indices for contingency tables under small sample sizes. *PLoS ONE*, 13, 8, e0199102, 1-19, 2018.

[91] Patriota, Alexandre Galvão. A classical measure of evidence for general null hypotheses. *Fuzzy Sets and Systems*, 233, 74-88, 2013.

[92] Patriota, Alexandre Galvão. On Some Assumptions of the Null Hypothesis Statistical Testing. *Educational and Psychological Measurement*, 77, 3, 507-528, 2017.

[93] Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. A Dynamic Software Certification and Verification Procedure. *Annals of the 5th ISAS-SCI*, 2, 426-435, 1999.

[94] Ranzato, Giulia; Ventura, Laura. *Biostatistica Bayesiana con "Matching Priors"*. Università degli Studi di Padova, Dipart. di Scienze Statistiche, 2018.

[95] Rifo, Laura Leticia Ramos; Torres, Soledad. Full Bayesian Analysis for a Class of Jump-Diffusion Models. *Communications in Statistics - Theory and Methods*, 38, 8, 1262-1271, 2009.

[96] Rifo, Laura Leticia Ramos; González-López, Veronica. Full Bayesian Analysis for a Model of Tail Dependence. *Communications in Statistics - Theory and Methods*, 41, 22, 4107-4123, 2012.

[97] Rincón, Sonia V. del; Rogers, Jeff; Widschwendter, Martin; Sun, Dahui; Sieburg, Hans B; Spruck, Charles. Development and Validation of a Method for Profiling Post-Translational Modification Activities Using Protein Microarrays. *PLoS ONE*, 5, 6, e11332, 1-11, 2010.

[98] Rodrigues, Josemar. Full Bayesian Significance Test for Zero-Inflated Distributions. *Communications in Statistics - Theory and Methods*, 35, 299-307, 2006.

[99] Santos, Natalia C.L; Dias, Rosa; Alvesc, Diego; Melo, Brian Ganassin, Maria; Gomes, Luiz; Severid, Willia; Agostinho, Angelo. Trophic and limnological changes in highly fragmented rivers predict the decreasing abundance of detritivorous fish. *Ecological Indicators*, 110, 105933, 1-8, 2020.

[100] Seixas, Andre Augusto Anderson; Hounie, Ana G; Fossaluza, Victor; Curi, Mariana; Alvarenga, Pedro G; Mathis, Maria A; Vallada, Homero; Pauls, David; Pereira, C.A.B; Miguel, Euripedes Constantino. Anxiety Disorders and Rheumatic Fever: Is There an Association? *CNS Spectrums*, 13, 12, 1039-1046, 2008.

[101] Shavitt, Roseli G; Requena, Guaraci; Alonso, Pino; Zai, Gwyneth; Costa, Daniel L.C; Pereira, C.A.B; Rosário, Maria C; Morais, Ivanil; Fontenelle, Leonardo; Cappi, C; Kennedy, James; Menchon, Jose M; Miguel, Eurípides; Richter, Peggy M.A. Quantifying dimensional severity of obsessive-compulsive disorder for neurobiological research. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 79, 206-212, 2017.

[102] Sikov, Anna; Stern, Julio M. Application of the full Bayesian significance test to model selection under informative sampling. *Statistical Papers*, 60, 89-104, 2019.

[103] Spektor, Mikhail S; Gluth, Sebastian; Fontanesi, Laura; Rieskamp, Jörg. How Similarity Between Choice Options Affects Decisions From Experience: The Accentuation-of-Differences Model. *Psychological Review*, 126, 1, 52-88, 2019.

[104] Stern, Julio Michael; Zacks, Shelemyahu. Testing the Independence of Poisson Variates under the Holgate Bivariate Distribution: The Power of a New Evidence Test. *Statistical and Probability Letters*, 60, 313-320, 2002.

[105] Takada, Hellinton Hatsuo; Ribeiro, Celma de Oliveira; Costa, Oswaldo Luiz do Valle; Stern, Julio Michael. Gini and Entropy-Based Spread Indexes for Primary Energy Consumption Efficiency and CO2 Emission. *Energies*, 13, 4938, 1-17, 2020.

[106] Vikas, K; Rao, Aruna K. Full Bayesian Empirical Likelihood Significance Test for Equality of Medians. *InterStat*, 2016, 01, 001, 1-9, 2016.

[107] Vosseler, Alexander; Weber, Enzo. Bayesian

analysis of periodic unit roots in the presence of a break. *Applied Economics*, 49, 38, 3841-3862, 2016.

[108] Wittenburg, Dörte; Teuscher, Friedrich; Klosa, Jan; Reinsch, Norbert. Covariance Between Genotypic Effects and its Use for Genomic Inference in Half-Sib Families. *G3 - Genes Genomes Genetics*, 6, 9, 2761-2772, 2016.

– General References in Logic and Statistics

[109] Aczél, János Dezsö; Daróczy, Zoltán. *On Measures of Information and Their Characterizations*. NY, Academic Press, 1975.

[110] Amari, Shun Ichi; Barndorff-Nielsen, Ole Eiler; Kass, Robert E; Lauritzen, Steffen L; Rao, Calyampudi Radhakrishna; Gupta, Shanti S. *Differential Geometry in Statistical Inference*. Hayward, IMS, 1987.

[111] Amari, Shun Ichi; Nagaoka, Hiroshi. *Methods of Information Geometry*. Providence, RI, AMS, 2007.

[112] Amari, Shun Ichi. *Information Geometry and Its Applications*. Berlin, Springer, 2016.

[113] Attneave, Fred. Applications of Information Theory to Psychology: A summary of basic concepts, methods, and results. NY, Holt, Rinehart and Winston, 1959.

[114] Barlow, Richard E; Prochan, Frank. *Statistical Theory of Reliability and Life Testing Probability Models*. Silver Spring, To Begin With, 1981.

[115] Basu, Debabrata; Ghosh, J.K. *Statistical Information and Likelihood*. Lecture Notes in Statistics, 45, 1988.

[116] Berger, James O; Wolpert, Rober L. *The Likelihood Principle*, 2nd ed. Hayward, CA, Inst of Mathematical Statistic, 1988.

[117] Bernardo, José M. Reference Analysis. p.17-90 in Dey, D.K; Rao, C.R. Bayesian Thinking: Modeling and Computation. *Handbook of Statistics*, v.25. Amsterdam, Elsevier, 2005.

[118] Béziau, Jean-Yves. The power of the hexagon. *Logica Universalis*, 6, 1-43, 2012.

[119] Birgin, Ernesto Julian Goldberg; Martínez, José Mario. *Practical Augmented Lagrangian Methods for Constrained Optimization*. SIAM, 2014.

[120] Birnbaum, Zygmunt Wilhelm; Esary, James D; Saunders, Samuel C. Multicomponent Systems and Structures, and their Reliability. *Technometrics*, 3, 55-77, 1961.

[121] Blanché, Robert. *Structures Intellectuelles: Essai sur l'Organisation Systématique des Concepts*. Vrin, 1966.

[122] Borges, Wagner de Souza. *Modelos Probabilísticos de Confiabilidade*. Rio de Janeiro, IMPA – Instituto de Matemática Pura e Aplicada, 1979.

[123] Box, George Edward Pelham; and Tiao, George C. *Bayesian Inference in Statistical Analysis*. London, Addison-Wesley, 1973.

[124] Bracewell, Ronald N. *The Fourier Transform and its Aplications*. McGraw-Hill, 1986.

[125] Brooks, Steve; Gelman, Andrew; Jones, Galin L; Meng, Xiao-Li. *Handbook of Markov Chain Monte Carlo*. CRC Press, 2011.

[126] Carnielli, Walter Alexandre; Pizzi, Claudio. *Modalities and Multimodalities*. Springer-Science, 2008.

[127] Carnielli, Walter Alexandre; Coniglio, Marcelo Esteban; Gabbay, Dov; Gouveia, P; Sernadas, C. *Analysis and Synthesis of Logics - How To Cut And Paste Reasoning Systems*. NY, Springer, 2008.

[128] Carnielli Walter Alexandre; Epstein, Richard L. *Pensamento Crítico*. São Paulo, Rideel, 2010.

[129] Chen, Ming-Hui; Shao, Qi-Man; Ibrahim, Joseph G. *Monte Carlo Methods in Bayesian Computation*. Springer-Verlag, 2000.

[130] Cherkassky, Vladimir; Mulier, Filip M. *Learning from Data*. NY, Wiley, 1998.

[131] Chernoff, Herman. On the Distribution of the Likelihood Ratio. *Ann. Math. Statist.* 25, 3, 573-578, 1954.

[132] Cohen, Jonathan. *The Probable and the Provable*. Oxford: Clarendon, 1977.

[133] Costa, Newton Carneiro da; Tsuji, Marcelo. Review of Theodore Hailperin, Sentential Probability Logic. *Modern Logic*, 8, 3-4, 103-107, 2001.

[134] Darwiche, Adnan Youssef; Ginsberg, Matthew L. A Symbolic Generalization of Probability Theory. *AAAI-92. 10-th Conf. American Association for Artificial Intelligence*, 1992.

[135] Darwiche, Adnan Youssef. *A Symbolic Generalization of Probability Theory*. Ph.D. Thesis, Stanford University, 1993.

[136] DeGroot, Morris Herman. Optimal Statistical Decisions. NY, McGraw-Hill, 1970.

[137] DeGroot, Morris H; Schervish, Mark J. *Probability and Statistics*. Pearson, 2011.

[138] Drton, Mathias. Likelihood Ratio Tests and Singularities. *Ann. Statist.*, 37, 2, 979-1012, 2009.

[139] Dubois, Didier; Prade, Henri. On Several Representations of an Uncertain Body of Evidence. pp.167-181 in Gupta, M.M; Sanchez, E; eds. *Fuzzy Information and Decision Processes*, North-Holland, 1982.

[140] Dubois,Didier; Foulloy, Laurent; Mauris, Gilles; Prade, Henri. Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10, 273–297, 2004.

[141] Dubois, Didier; Prade, Henri. Rough Fuzzy Sets and Fuzzy Rough Sets. *International Journal of General Systems*, 2007. doi:10.1080/03081079008935107

[142] Dudois, Didier; Prade, Henri. From Blanché's Hexagonal Organization of Concepts to Formal Concept Analysis and Possibility Theory. *Logica Universalis*, 6, 149-169, 2012.

[143] Dugdale, J. Sydney. Entropy and Its Physical Meaning. London, Taylor and Francis, 1996.

[144] Dugundji, James. *Topology*. Boston, Allyn and Bacon, 1966.

[145] Eaton, Morris L. *Group Invariance Applications in Statistics*. Hayward, CA, IMS, 1989.

[146] Ebbinghaus, Heinz-Dieter; Flum, Jörg. *Finite Model Theory*. Berlin, Springer, 1999.

[147] Epstein, George; Frieder, Guideon; Rine, David. The Development of Multiple-Valued Logic as Related to Computer Science. *Computer*, 7, 9, 20-32, 1974.

[148] Epstein, George. *Multiple-Valued Logic Design*. IOP Publishing, 1993.

[149] Epstein, Richard L; Carnielli Walter Alexandre. *Computability: Computable Functions, Logic, and the Foundations of Mathematics*. Socorro, NM, Advanced Reasoning Forum, 2008.

[150] Evans, Michael. Bayesian Inference Procedures Derived via the Concept of Relative Surprise. *Communications in Statistics*, 26, 1125-1143, 1997.

[151] Fang, Shu Cheng; Rajasekera, Jay R; Tsao, Jacob, H.S. *Entropy Optimization and Mathematical Programming*. Kluwer, Dordrecht, 1997.

[152] Finetti, Bruno de. *Matematica Logico-Intuitiva*. Roma, Cremonese, 1957.

[153] Finetti, Bruno de. *Theory of Probability*. New York, Wiley, 1997.

[154] Flanders, Harley; Korfhage, RobertR; Price, Justin J. *A Second Curse in Calculus*. NY, Academic Press, 1974.

[155] Fong, Brendan; Spivak, David Isaac. *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*. Cambridge Univ. Press, 2019

[156] Fossaluza, Victor; Lauretto, Marcelo de Souza; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. Combining Optimization and Randomization Approaches for the Design of Clinical Trials. *Springer Proceedings in Mathematics and Statistics*, 118, 173-184, 2015.

[157] Frieden, Roy Bernard *Science from Fisher Information: A Unification*. Cambridge Univ. Press, 2004.

[158] Gaifman, Haim. Concerning Measures in First Order Calculi. *Israel J. of Mathematics*, 2, 1–18, 1964.

[159] Gaifman, Haim; Snir, Marc. Probabilities Over Rich Languages, Testing and Randomness. *The Journal of Symbolic Logic*, 47, 3, 495-548, 1982.

[160] Gelman, Andrew; Carlin, John B; Stern, Hal S; Rubin, Donald B. *Bayesian Data Analysis*. Chapman-Hall/ CRC, 2004.

[161] Gilks, Wally R; Richardson, Sylvia; Spiegelhalter, David John. *Markov Chain Monte Carlo in Practice*. CRC Press, 1996.

[162] Gottenwald, Siegfried. *A Treatise on Many-Valued Logics*. Research Studies Press, 2001.

[163] Guy, Dan M; Carmichael, Douglas; Whittington, O. Ray. *Audit Sampling: An Introduction*. John Wiley, 1994.

[164] Hailperin, Theodore. *Sentential Probability Logic: Origins, Development, Current Status, and Technical Applications*. Bethlehem, Lehigh Univ. Press, 1996.

[165] Hailperin, Theodore. *Logic with a Probability Semantics: Including Solutions to Some Philosophical Problems*. Bethlehem, Lehigh Univ. Press, 2011.

[166] Hammersley, John Michael; Handscomb, David Christopher. *Monte Carlo Methods*. Chapman and Hall, 1964.

[167] Hedman, Shawn. *A First Course in Logic: An Introduction to Model Theory, Proof Theory, Computability, and Complexity*. Oxford Univ. Press, 2004.

[168] Hocking, Ronald R. The Analysis of Linear Models. Monterey: Brooks Cole, 1985.

[169] Howson, Colin. Can Logic be Combined with Probability? Probably. *Journal of Applied Logic*, 7, 2, 177–187, 2009.

[170] Hutter, Marcus. *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Berlin, Springer, 2005.

[171] Hutter, Marcus. *Predictive Hypothesis Identification*. Canberra ACT, 2008. arXiv:0809.1270v1

[172] Inhasz, Rafael; Stern, Julio Michael. Emergent Semiotics in Genetic Programming and the Self-Adaptive Semantic Crossover. *Studies in Computational Intelligence*, 314, 381-392, 2010.

[173] Jeffreys, Harold. *Theory of Probability*. Oxford, Clarendon Press. 1961, first ed. 1939.

[174] Kadane, Joseph Born. Pragmatics of Uncertainty. NY, Chapman-Hall/CRC, 2016.

[175] Kaplan, Stanley; Lin, James C. An Improved Condensation Procedure in Discrete Probability Distribution Calculations. *Risk Analysis*, 7, 15-19, 1987.

[176] Kapur, Jagat Narain. *Maximum Entropy Models in Science and Engineering*. New Delhi, John Wiley, 1989.

[177] Kapur, Jagat Narain; Kesavan, Hiremagalur Krishnaswamy. *Entropy Optimization and Applications*. Academic Press, 1992

[178] Karawatzki, Roman; Leydold, Josef; Pötzelberger, Klaus *Automatic Markov Chain Monte Carlo Procedures for Sampling from Multivariate Distributions*. `epub.wu.ac.at/id/eprint/1400`

[179] Kaufmann, Arnold; Grouchko, Daniel; Cruon, R. *Mathematical Models for the Study of the Reliability of Systems*. NY, Academic Press, 1977.

[180] Klir, George J; Folger, Tina. *Fuzzy Sets, Uncertainty, and Information*. Prentice-Hall, 1988.

[181] Klir, George J; Yuan, Bo. *Fuzzy Sets and Fuzzy Logic*. Prentice-Hall, 1995.

[182] Lauretto, Marcelo de Souza; Nakano, Fabio; Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. Intentional sampling by goal optimization with decoupling by stochastic perturbation. *AIP Conference Proceedings*, 1490, 189-201, 2012.

[183] Lawvere, F. William; Schanuel, Stephen Hoel. *Conceptual Mathematics: A First Introduction to Categories*. Cambridge Univ. Press, 1997.

[184] Lindley, Dennis Victor. *Bayesian Statistics: A Review*. SIAM, 1972.

[185] Lindley, Dennis Victor. Subjective Probability, Decision Analysis and their Legal Consequences. *J. R. Stat. Soc.*, 154, 83-92.

[186] Lipschutz, Seymour. *General Topology*. NY, McGraw-Hill, 1965.

[187] Luenberger, David Gilbert; Ye, Yinyu. *Linear and Nonlinear Programming*. NY, Springer, 2008.

[188] Minoux, Michel. *Mathematical Programming: Theory and Algorithms*. John Willey, 1986.

[189] Pawitan Yudi. *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press, 2001.

[190] Pereira, Carlos Alberto de Bragança; Nakano, Fabio; Stern, Julio Michael. Actuarial Analysis via Branching Processes. *Annals of the 6th ISAS-SCI*, 8, 353-358, 2000.

[191] Pereira, Basilio de Bragança; Pereira, Carlos Alberto de Bragança. A Likelihood Aproach to Diagnostic Tests in Clinical Medicine. *RevStat - Statistical Journal*, 3, 1, 77-98, 2005.

[192] Pereira, Carlos Alberto de Bragança; Stern, Julio Michael. Special Characterizations of Standard Discrete Models. *RevStat - Statistical Journal*, 6, 199-230, 2008.

[193] Pereira, Basilio de Bragança; Pereira, Carlos Alberto de Bragança. *Model Choice in Nonnested Families*. Berlin, Springer. 2016.

[194] Pflug, Georg Christian. *Optimization of Stochastic Models: The Interface Between Simulation and Optimization*. Dordrecht, Kluwer, 1996.

[195] Piskunov, Nikilai. *Differential and Integral Calculus*. Moscow, MIR, 1969.

[196] Priest, Graham. *An Introduction to Non-Classical Logics*. Cambridge Univ. Press, 2001.

[197] Rine, David C. *Computer Science and Multiple-Valued Logic*. North-Holland, 1984.

[198] Ripley, Brian David. Stochastic Simulation. NY, Wiley, 1987.

[199] Rissanen, Jorma. *Stochastic Complexity in Statistical Inquire*. Singapore, World Scientific, 1989.

[200] Rodrigues, Abilio; Bueno-Soler, Juliana; Carnielli, Walter Alexandre. Measuring Evidence: A Probabilistic Approach to an Extension of Belnap-Dunn Logic. *Synthese*, 198, 22, 5451-5480, 2021.

[201] Romero, Carlos. *Handbook of Critical Issues in Goal Programming*. Oxford, Pergamon Press, 1991.

[202] Royall, Richard. *Statistical Evidence: A Likelihood Paradigm*. London, Chapman and Hall. 1997.

[203] Ruhl, John B. Law's Complexity: A Primer. *Georgia State Univ. Law Review*, 24, 4, 1-28, 2008.

[204] Ruhl, John B; Katz, Daniel Martin. Measuring, Monitoring, and Managing Legal Complexity. *University of Iowa Law Review*, 101, 191-244, 2015.

[205] Russell, Stuart. Unifying Logic and Probability. *Communications of the ACM*, 58, 7, 88-97, 2015.

[206] Saa, Olivia Terence; Stern, Julio Michael. Auditable blockchain randomization tool. *Proceedings*, 33, 1, 17.1-17.6, 2019.

[207] Sakamoto,Yosiyuki; Ishiguro, Makio; Kitagawa, Genshiro. *Akaike Information Criterion Statistics*. Dordrecht, Kluwer, 1986.

[208] Salicone, Simona. *Measurement Uncertainty: An Approach via the Mathematical Theory of Evidence*. Springer, 2007.

[209] Schwartzman, Armin; Mascarenhas, Walter F; Taylor, Jonathan E. Inference for Eigenvalues and Eigenvectors of Gaussian Symmetric Matrices. *Ann. Statist.*, 36, 6, 2886-2919, 2008.

[210] Searcóid, Mícheál Ó. *Metric Spaces*. London, Springer, 2006.

[211] Self, Steven G; Liang, Kung-Yee. Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions. *J. Amer. Statist. Assoc.*, 82, 398, 605-610, 1987.

[212] Shackle, George Lennox Sharman. *Uncertainty in Economics and Other Reflections*. London, Cambridge Univ. Press, 1968.

[213] Shackle, George Lennox Sharman. *Decision, Order and Time in Human Affairs*. London, Cambridge

Univ. Press, 1969.

[214] Shier, Douglas R. *Network Reliability and Algebraic Structures*. Oxford, Clarendon Press, 1991.

[215] Sichelman, Ted. Quantifying Legal Entropy. *Frontiers in Physics*, 21, 9, 665054, 1-14, 2021.

[216] Skilling, John. Nested Sampling. *AIP Conference Proceedings*, 735, 395-405, 2004.

[217] Skilling, John. Nested Sampling for General Bayesian Computation. *Bayesian Analysis*, 1, 4, 833-860, 2006.

[218] Spall, James C. *Introduction to Stochastic Search and Optimization*. Willy, 2003.

[219] Spivak, David Isaac. *Category Theory for the Sciences*. MIT Press, 2014.

[220] Springer, Melvin D. *The Algebra of Random Variables*. NY, Wiley, 1979.

[221] Stern, Julio Michael. Simulated Annealing with a Temperature Dependent Penalty Function. *ORSA Journal on Computing*, 4, 311-319, 1992.

[222] Stern, Julio Michael; Colla, Ernesto Coutinho. Factorization of Bayesian Networks. pp.275-294 in Nakamatsu, K; Phillips-Wren, G; Jain, L.C; Howlett, R.J. eds. *New Advances in Intelligent Decision Technologies*. Heidelberg, Springer, 2009.

[223] Thulin, Mans. Decision-theoretic justifications for Bayesian hypothesis testing using credible sets. *Journal of Statistical Planning and Inference*, 146, 133-138, 2014.

[224] Vieland, Veronica J; Chang, Hasok. No evidence amalgamation without evidence measurement. *Synthese*. 196, 3139-3161, 2019.

[225] Voigt, Hans-Michael; Mühlenbein, Henz; Schwefel, Hans-Paul. *Evolution and Optimization '89*. Berlin, Akademie Verlag, 1990.

[226] Wah, Benjamin W; Chen, Yixin; Wang, Tao. Theory and Applications of Simulated Annealing for Nonlinear Constrained Optimization. Ch.9, pp.155-186 in Tan, Cher Ming. *Simulated Annealing*, 2008.

[227] Wallace, Christopher Stewart; Boulton David Morris An Information Measure for Classification. *Computer Journal*, 11, 2, 185-194, 1968.

[228] Wallace, Christopher Stewart; Dowe, David. Minimum Message Length and Kolmogorov Complexity. *Computer Journal*, 42, 4, 270-283, 1999.

[229] Wallace, Christopher Stewart. *Statistical and Inductive Inference by Minimum Message Lenght*. NY, Springer, 2005.

[230] Wechsler, Sergio; Pereira, Carlos Alberto de Bragança; Marques, Paulo Cilas. Birnbaum's Theorem Redux. *AIP Conf. Proceedings*, 1073, 96-100, 2008.

[231] West, Mike; Harrison, Jeff. *Bayesian Forecasting and Dynamic Models*. Springer, 1997.

[232] Williams, David. Weighing the Odds. Cambridge University Press, 2001.

[233] Williamson, Robert Charles. *Probabilistic Arithmetic*. Ph.D. Thesis, Univ. of Queensland, 1989.

[234] Zeleny, Milan. *Multiple Criteria Decision Making*. NY, McGraw-Hill, 1982.

[235] Zellner, Arnold. *Introduction to Bayesian Inference in Econometrics*. NY, Wiley, 1971.

– General References in Law and Philosophy

[236] Barker, Frank; Evans Robin. *The Probability of Mr Bayes: A constructive re-evaluation of Mr Bayes' essay and of the opinions concerning it expressed by various authorities*. Technical report, Department of Electrical and Electronic Engineering, University of Melbourne, 2001.

[237] Barak, Aharon. *Proportionality: Constitutional Rights and Their Limitations*. Cambridge Univ. Press, 2012.

[238] Bayes, Thomas. An Essay Towards Solving a Problem in the Doctrine of Chances. *Phil. Trans. Roy. Soc. London*, 53, 370-418, 1764.

[239] Bueno-Soler, Juliana; Carnielli, Walter Alexandre. Paraconsistent Probabilities: Consistency, Contradictions and Bayes' Theorem. *Entropy*, 18, 325, 1-18, 2016.

[240] Campilongo, Celso Fernandes; Amato, Lucas Fucci; Barros, Marco Antonio Loschiavo Leme de. *Luhmann and Socio-Legal Research: An Empirical Agenda for Social Systems*. NY, Routledge, 2020.

[241] Cauble, Emily. Unsophisticated Taxpayers, Rules Versus Standards, and Form Versus Substance. *Loyola University Chicago Law Journal*, 52, 329-349, 2021.

[242] Condé, Mauro Lúcio Leitão. *Wittgenstein, Linguagem e Mundo*. São Paulo, Annablume, 1998.

[243] Dumitriu, Anton. *History of Logic*. Abacus Press, 1977.

[244] Engle, Eric Allen. The History of the General Principle of Proportionality. *Dartmouth Law Journal*, 10, 1, 1–11, 2012.

[245] Finetti, Bruno de. *L'invenzione della verità*. Milano, Scienza e Idee, 2006.

[246] - Foerster, Heinz von. *Understanding Understanding: Essays on Cybernetics and Cognition*. NY, Springer Verlag, 2003.

[247] Foerster, Heinz von; Segal, Lynn. The Dream of Reality. Heintz von Foerster's Constructivism. NY, Springer, 2001.

[248] Gallais, Pierre. *Dialectique du Récit Mediéval: Chrétien de Troyes et l'Hexagone Logique*. Amster-

dam, Rodopi, 1982.

[249] Garson, James W. *Modal Logic for Philosophers*. Cambridge Univ. Press, 2006.

[250] Gaskins, Richard H. Burdens of Proof in Modern Discourse. New Haven, Yale Univ. Press, 1992.

[251] Good, Irving John. *Good Thinking*. Univ. of Minnesota. 1983.

[252] Hulsroj, Peter. *The Principle of Proportionality*. Springer Netherlands, 2013.

[253] Kades, Eric. *The Laws of Complexity and the Complexity of Laws: The Implications of Computational Complexity Theory for the Law*. College of William and Mary Law School Scholarship Repository, 1997.

[254] Kokott, Juliane. *The Burden of Proof in Comparative and International Human Rights Law*. The Hague: Kluwer, 1998.

[255] Luhmann, Niklas. *A Sociological Theory of Law*. London, Routledge & Kegan Paul, 1985.

[256] Luhmann, Niklas. Ecological Communication. Chicago Univ. Press, 1989.

[257] Maturana, Humberto Romesín; Varela, Francisco Javier. *Autopoiesis and Cognition. The Realization of the Living*. Dordrecht, Reidel, 1980.

[258] Muchmore, Adam I. Uncertainty, Complexity, and Regulatory Design. *Houston Law Review*, 53, 5, 1321-1367, 2016.

[259] Murray, Jamine; Webb, Thomas E; Wheatley, Steven. *Complexity Theory and Law: Mapping an Emergent Jurisprudence*. London, Routledge, 2019.

[260] Pigliucci, Massimo; Boudry, Maarten. Prove it! The Burden of Proof Game in Science vs. Pseudoscience Disputes. *Philosophia*, 42, 2, 487-502, 2014.

[261] Rasch, William. *Niklas Luhmann's Modernity: Paradoxes of Differentiation*. Stanford Univ. Press, 2000.

[262] Schuck, Peter H. Legal Complexity: Some Causes, Consequences, and Cures. *Duke Law Journal*, 42, 1, 1-52, 1992.

[263] Teubner, Gunther. *Autopoietic Law: A New Approach to Law and Society*. Berlin, Walter de Gruyer, 1988.

[264] Wittgenstein, Ludwig. Logisch-Philosophische Abhandlung, *Annalen der Naturphilosophie*, 14, 185-262, 1921. Translated as *Tractatus Logico Philosophicus*, NY, Harcourt & Brace, 1922.

[265] Zeleny, Milan. *Autopoiesis, Dissipative Structures and Spontaneous Social Order*. Boulder, CO, AAAS, 1979.